



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS GRADUAÇÃO EM QUÍMICA

**SIMULAÇÃO COMPUTACIONAL DAS INTERAÇÕES ENTRE O HERBICIDA
GLIFOSATO E A ENZIMA 5-ENOLPIRUVIL-SHIKIMATO-3-FOSFATO SINTASE
(EPSPs) NO DESENVOLVIMENTO DE UM NANOBIOSSENSOR**

Moacir Fernandes Ferreira Júnior

UBERLÂNDIA-MG

2016

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS GRADUAÇÃO EM QUÍMICA
DOUTORADO

**SIMULAÇÃO COMPUTACIONAL DAS INTERAÇÕES ENTRE O HERBICIDA
GLIFOSATO E A ENZIMA 5-ENOLPIRUVIL-SHIKIMATO-3-FOSFATO SINTASE
(EPSPs) NO DESENVOLVIMENTO DE UM NANOBIOSSENSOR**

Tese apresentada como parte dos
requisitos para obtenção do título de
DOUTOR EM QUÍMICA do programa
de pós-graduação em Química do IQ-
UFU.

Orientador: Prof. Dr. Eduardo de Faria Franca

UBERLÂNDIA-MG

2016

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

F383
2016

Ferreira Júnior, Moacir Fernandes, 1983-
Simulação computacional das interações entre o herbicida Glifosato e a enzima 5-Enolpiruvil-Shikimato-3-Fosfato Sintase (EPSPs) no desenvolvimento de um nanobiossensor / Moacir Fernandes Ferreira Júnior. - 2016.
102 f. : il.

Orientador: Eduardo de Faria Franca.
Tese (doutorado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Química.
Inclui bibliografia.

1. Química - Teses. 2. Dinâmica molecular - Teses. 3. Herbicidas - Teses. 4. Microscopia de força atômica - Teses. I. Franca, Eduardo de Faria. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Química. III. Título.



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA



**Ata da defesa de TESE DE DOUTORADO junto ao Programa de Pós-Graduação em Química,
do Instituto de Química da Universidade Federal de Uberlândia
DEFESA DE TESE DE DOUTORADO EM QUÍMICA, NÚMERO 57/PPQUI.**

DATA: 21/12/2016

DISCENTE: Moacir Fernandes Ferreira Júnior

MATRÍCULA: 11213QMI008

TÍTULO DO TRABALHO: Simulação computacional das propriedades e interações entre o herbicida glifosato e a enzima enol-pirúvil shiquimato fosfato sintase (EPsPs) no desenvolvimento de um nanobiossensor.

ÁREA DE CONCENTRAÇÃO: Química

LINHA DE PESQUISA: Físico-Química de macromoléculas e coloides

PROJETO DE PESQUISA DE VINCULAÇÃO: Controle molecular no desenvolvimento de nanobiossensores.

Às oito horas do dia vinte e um de dezembro do ano dois mil e dezesseis, no Auditório Prof. Manuel Gonzalo Hernández Terrones, Bloco 5I, Campus Santa Mônica, reuniu-se a Banca Examinadora composta pelos Professores Doutores Odonório Abrahão Júnior, da Universidade Federal do Triângulo Mineiro, Osmair Vital de Oliveira, do Instituto Federal de São Paulo, Teodorico de Castro Ramalho, da Universidade Federal de Lavras, por meio de videoconferência, Carlos Alberto de Oliveira, do Instituto de Química da UFU e Eduardo de Faria Franca, orientador(a) e presidente da mesa. Iniciando os trabalhos, o(a) presidente da mesa apresentou o(a) candidato(a) e a Banca Examinadora, agradeceu a presença do público e discorreu sobre as normas e critérios para a realização desta sessão, baseadas na Norma Regulamentar nº 03/2013/PPQUI. Em seguida, o(a) presidente da mesa concedeu a palavra ao(à) candidato(a) para a exposição do seu trabalho e, em sequência, aos examinadores, em ordem sucessiva, para arguir o(a) apresentador(a). A duração da apresentação e o tempo de arguição e resposta deram-se conforme as normas do Programa. Ultimada a arguição, desenvolvida dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu os conceitos finais e A provou o(a) candidata(o). Por sugestão da Banca Examinadora, o título do trabalho será

SIMULAÇÃO COMPUTACIONAL DAS INTERAÇÕES ENTRE
O HERBICIDA GLIFOSATO E A ENZIMA 5-ENOLPIRUVIL-
SHIKIMATO-3-FOSFATO SINTASE (EPsPs) NO
DESENVOLVIMENTO DE UM NANOBIOSSENSOR

Esta defesa de Tese de Doutorado é parte dos requisitos necessários à obtenção do título de Doutor. O competente diploma será expedido após cumprimento do estabelecido nas normas do Programa, legislação e regulamentação internas da UFU. As correções observadas pelos examinadores deverão ser realizadas no prazo máximo de 45 dias. Nada mais havendo a tratar, deu-se por encerrada a sessão às 18 horas e 08 minutos e lavrada a presente ata, que após lida e aprovada, foi assinada pela Banca Examinadora.

Prof. Dr. Odonório Abrahão Júnior – UFTM

Prof. Dr. Osmair Vital de Oliveira – IFSP

Prof. Dr. Teodorico de Castro Ramalho –
UFLA

Prof. Dr. Carlos Alberto de Oliveira – IQUFU

Professor(a) Orientador(a): Eduardo de Faria Franca – IQUFU

AGRADECIMENTOS

- Em especial ao meu orientador Eduardo de Faria Franca, pela dedicação na minha formação como aluno e pela amizade adquirida nesse período.
- À minha família, esposa, pais, irmã e sobrinhos pelo carinho, paciência e incentivo nos momentos de maior necessidade.
- Aos amigos que fizeram parte desses momentos sempre me ajudando e incentivando.
- Ao meu coorientador Fábio de Lima Leite.
- Aos meus colegas e amigos de laboratório que participaram direta e indiretamente deste trabalho.

“Nós não nos perguntamos para que propósito os pássaros cantam, pois a canção é o seu prazer, já que foram criados para cantar. E do mesmo modo nós não devemos nos perguntar porque a mente humana se dá ao trabalho de sondar os segredos dos céus... A diversidade de fenômenos da natureza é tão grande e os tesouros ocultos nos céus tão ricos precisamente para que a mente humana nunca careça de seu alimento”.

Johannes Kepler

LISTA DE FIGURAS

Figura 1 - Estrutura e classificação dos aminoácidos.....	15
Figura 2 - Ligação peptídica.....	16
Figura 3 - Cadeia peptídica: Em cinza a cadeia principal, formada através das ligações peptídicas, e em vermelho as cadeias laterais.....	17
Figura 4 - Grupo peptídico planar: (a) Caráter de dupla ligação da ligação peptídica;(b) As ligações N—C α e C α —C podem sofrer rotação em ângulos denominados ϕ e ψ , respectivamente.....	18
Figura 5 - Diagrama de Ramachandran para resíduos de L-Ala. As conformações dos peptídeos são definidas pelos valores de ψ e ϕ . As áreas em azul-escuro refletem conformações que não envolvem sobreposições estéricas, permitidas; o azul-médio indica conformações permitidas nos limites extremos para os contatos atômicos desfavoráveis; a área azul-clara reflete conformações que são permitidas se houver uma pequena flexibilidade nos ângulos de ligação e área em cinza conformações proibidas.....	19
Figura 6 - Estruturas em (a) α -hélice e (b) folha β pregueada.....	20
Figura 7 - Conversão do shikimato a corismato na rota shikimato.....	22
Figura 8 - Representação (em cartoon) da conformação da enzima EPSPs; (a) Estado não ligado (aberto) e (b) Estado ligado (fechado). O S3P e o glifosato são representados nos modelos (bola e bastão) e (van der Waals-VDW), respectivamente.....	23
Figura 9 - Fração de ocorrência das espécies de glifosato construído a partir dos valores de pKa (marrom, composto com carga formal igual a 0; verde, composto com carga formal igual a -1; azul, composto com carga formal igual a -2; vermelho composto com carga formal igual a -3 ; e representação dos equilíbrios.....	24
Figura 10 - Rotas de decomposição microbiológica do glifosato.....	25
Figura 11 - (a) Ilustração do princípio de funcionamento do AFM; (b) Microscopia eletrônica de varredura de um cantiléver de nitreto de silício com agulha acoplada (em detalhe).....	27
Figura 12 - Curva de força pela distância de separação entre a agulha e amostra.....	28
Figura 13 - Representação esquemática dos modos de operação em AFM: (a) modo contato, (b) modo não-contato e (c) modo intermitente.....	29
Figura 14 - Esquema representativo para o sistema ponta funcionalizada-substrato em um AFM.....	30
Figura 15 - Ângulo diedro.....	40
Figura 16 - Representação das Condições Periódicas de Contorno em três dimensões.....	51
Figura 17 - Representação do raio de corte esférico em duas dimensões.....	52

Figura 18 - Construção de um sistema de células periódicas no método de Ewald.....	52
Figura 19 - Método do campo de reação. A seta com a haste sombreada representa a soma dos dipolos das moléculas dentro da esfera de raio de corte r_c	56
Figura 20 - Representação dos modelos rígidos na solvatação explícita; (a) de três sítios de interação – SPC, TIP3P; (b) de quatro sítios – TIP4P; (c) de cinco sítios – TIP5P; l_1 e l_2 representam as distâncias entre o átomo hidrogênio e oxigênio; q_1 e q_2 representam as cargas atômicas do hidrogênio e oxigênio, respectivamente; θ e Φ são os ângulos de ligação e σ é a distância para o potencial zero.....	58
Figura 21 - Representação esquemática da dissociação do herbicida glifosato, em $pK_a = 5,69$	63
Figura 22 - Representação da molécula glifosato: Diedros parametrizados HO3-O3-P-C1 e P-C1-N-HN1.....	63
Figura 23 - Diagrama de Ramachandran indicando as distribuições dos ângulos torcionais Φ e ψ da cadeia principal. As regiões mais favoráveis são mostradas em vermelho, as adicionalmente permitidas em marrom, as generosamente permitidas em amarelo e as não permitidas em amarelo claro.....	70
Figura 24 - Representação da sobreposição da estrutura da enzima EPSPs (representação new cartoon) , antes da simulação por dinâmica molecular (em cinza) e após a simulação (em verde).....	72
Figura 25 - (A) Raiz quadrada do desvio quadrático médio - RMSD e (B) raio de giro - RG da enzima EPSPs.....	73
Figura 26 - Potencial eletrostático plotado na superfície molecular da enzima EPSPs. Superfície no eixo -x voltado para frente (A); superfície no eixo y voltado para frente (B) e superfície no eixo -y voltado para frente (C).....	74
Figura 27 - Representação da sobreposição da molécula glifosato (representada em CPK) antes e após a simulação por dinâmica molecular.....	76
Figura 28 - Raiz quadrada do desvio quadrático médio (RMSD) do herbicida glifosato.....	76
Figura 29 - Perfil de energia calculados com cálculo quântico (vermelho) e cálculo clássico usando o método dos mínimos quadrados (preto) para os diedros HO3-O3-P-C1 e P-C1-N-HN1.....	78
Figura 30 - Distribuição do ângulo diedral, para o diedro HO3-O3-P-C1 (acima) e para o diedro P-C-N-HN1 (abaixo).....	79
Figura 31 - Representação da sobreposição do sistema EPSPs-S3P-GPJ. Enzima (representada em newcartoon), shikimato-3-fosfato (em CPK) e o herbicida glifosato (representado em	

superfície de van der Waals). EPSPs antes da simulação por dinâmica molecular (em cinza) e após a dinâmica molecular (em verde).....	80
Figura 32 - Raiz quadrada do desvio quadrático médio (RMSD) da EPSPs na presença do substrato S3P e do inibidor GPJ (preto) e da EPSPs isolada (vermelho).....	81
Figura 33 - (A) Número de ligações de hidrogênio entre a enzima sozinha (vermelho) e a água e entre a enzima com S3P e água (preto) e (B) interações intermoleculares entre os resíduos do sítio ativo na enzima sozinha (vermelho) e na enzima com S3P (preto).....	82
Figura 34 - Flutuação quadrática média (RMSF) de 8 resíduos de aminoácidos do sítio ativo na presença do S3P-GPJ (preto) e na enzima isolada (vermelho).....	83
Figura 35 - Dinâmica Molecular Direcional em três possíveis caminhos de retirada do herbicida glifosato. (A) Ao longo do eixo x; (B) Ao longo do eixo y e (C) ao longo do eixo -y.	85
Figura 36 - Perfil de força aplicado ao longo dos eixos x (preto), y (azul) e -y (vermelho).....	86
Figura 37 - Distribuição radial de pares nos três eixos: Eixo x (preto); Eixo y (azul) e Eixo -y (vermelho).....	88
Figura 38 - Variação do RMSD da enzima EPSPs nas simulações de dinâmica molecular direcional ao longo dos eixos x (preto), y (azul) e -y (vermelho).....	89
Figura 39 - Gráfico Quantil-Quantil (I) – Eixo x, (a) trabalho a 300ps, (b) trabalho a 500ps, (c) trabalho a 1000ps, (d) trabalho a 1500ps; (II) – Eixo y, (a) trabalho a 500ps, (b) trabalho a 1000ps, (c) trabalho a 1500ps, (d) trabalho a 2000ps; (III) – Eixo -y, (a) trabalho a 300ps, (b) trabalho a 1000ps, (c) trabalho a 1500ps, (d) trabalho a 1750ps;.....	90
Figura 40 - Potencial de Força Média das simulações de SMD ao longo dos três eixos: x (em preto); y(em azul) e -y (em vermelho).....	92

LISTA DE TABELAS

Tabela 1 - Parâmetros de estiramento de ligação, flexão angular e diedral empregados no campo de força OPLS-AA.....	64
Tabela 2 - Estatística da estereoquímica do modelo 3D da enzima EPSPs.....	71
Tabela 3 - Carga parcial dos átomos do glifosato.....	75
Tabela 4 - Coeficientes e ângulos dos diedros parametrizados.....	77
Tabela 5 - Energia livre de solvatação da enzima EPSPs com o substrato S3P e da enzima sozinha.....	81
Tabela 6 - Teste de Shapiro-Wilk.....	91

LISTA DE ABREVIATURAS E SIGLAS

AFM	Atomic Force Microscope
AM1	Austin Model 1
AMPA	Ácido Aminometilfosfônico
ANVISA	Agência Nacional de Vigilância Sanitária
CFM	Chemical Force Microscopy
DFT	Density Functional Theory
EPSPs	5-Enolpiruvil-Shikimato-3-Fosfato Sintase
ES	Enzima-Substrato
GPJ	Glifosato
HF	Hartree-Fock
Iarg	Agência Internacional de Pesquisa sobre o Câncer
MD	Molecular Dynamics
MM	Mecânica Molecular
MNDO	Modified Neglect of Diatomic Overlap
OMS	Organização Mundial da Saúde
RCSC-PDB	Research Collaboratory for Structural Bioinformatics – Protein Data Bank
PEP	Fosfoenol-Piruvato
PM3	Parametric Method 3
PM6	Parametric Method 6 Hamiltonian
PM7	Parameterized Model number 7
PME	Particle Mesh Ewald
PMF	Potential of Mean Force
RESP	Restrained Electrostatic Potential
RG	Raio de Giro
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
S3P	Shikimato-3-Fosfato
SCF	Self-Consistent-Field
SMD	Steered Molecular Dynamics
SPC	Simple Point Charge
TIP	Transferable Intermolecular Potentials
@	Marca Registrada
CPU	Central Processing Unit

RESUMO

Este trabalho teve como objetivo, desenvolver um modelo molecular que possa descrever propriedades de equilíbrio, bem como as forças de interação envolvidas em um sistema enzima-inibidor, onde a enzima é a 5-Enolpiruvil-Shikimato-3-Fosfato Sintase (EPSPs) e seu inibidor o herbicida glifosato (GPJ). A participação do substrato Shikimato-3-Fosfato (S3P) também foi estudada. A estabilidade e a distribuição de cargas ao longo da superfície da enzima EPSPs foram analisadas para subsidiar o desenvolvimento de um nanobiossensor específico para a detecção do herbicida glifosato através da funcionalização da ponta de um microscópio de força atômica (AFM) com a enzima EPSPs. A parametrização de dois ângulos diedros para o herbicida GPJ foram obtidos através de cálculos mecânico-quânticos ao nível de Hartree Fock, utilizando como conjunto de bases o 6-31G*. Os parâmetros obtidos, relacionados aos coeficientes de Ryckaert-Bellemans, foram implementados no campo de força OPLS-AA, utilizado pelo programa GROMACS. Simulações de dinâmica molecular direcional (SMD), foram realizadas em três possíveis caminhos de retirada do herbicida do sítio ativo, aplicando-se forças harmônicas externas com constante de força equivalente à do cantiléver do AFM, com o objetivo de simular experimentos de detecção do herbicida por microscopia de força atômica. Uma expansão cumulante de segunda ordem da igualdade de Jarzynski foi utilizada para os cálculos do perfil de energia livre ao longo da coordenada de reação e, portanto, para se obter o Potencial de Força Média (PMF) das simulações de SMD. Os três caminhos escolhidos apresentaram diferentes barreiras energéticas relacionadas com o rompimento das interações do inibidor com a enzima. Estas barreiras energéticas calculadas pelas simulações estão na mesma ordem de magnitude das energias livres de ligação para sistemas semelhantes obtidas a partir das constantes de inibição experimentais.

Palavras-chave: Dinâmica Molecular, Enzima EPSPs, Glifosato, AFM, Potencial de Força Média.

ABSTRACT

This work aims to develop a molecular model that can describe equilibrium properties, as well as the interacting forces involved in an enzyme-inhibitor system, where the enzyme is the 5-Enolpyruvyl-Shikimate-3-Phosphate Synthase (EPSPs) and its inhibitor is the glyphosate herbicide (GPJ). The participation of the substrate Shikimate-3-Phosphate (S3P) was also studied. The stability and distribution of charges along the surface of the EPSPs enzyme were analyzed for the development of a specific nanobiosensor for glyphosate herbicide detection through the functionalization of the tip of an atomic force microscope (AFM) with the enzyme EPSPs. The parameterization of two dihedral angles for the GPJ herbicide was obtained through mechanical-quantum calculations at the Hartree Fock level using the 6-31G* base set. The parameters obtained, related to the Ryckaert-Bellemans coefficients, were implemented in the OPLS-AA force field, used by the GROMACS program. Steered Molecular Dynamics Simulations (SMD) were performed in three possible unbinding pathways, to remove the herbicide from the active site, applying external harmonic forces with a force constant equivalent to that of the AFM cantilever, in order to simulate experiments to detect the herbicide by atomic force microscopy. A second-order cumulant expansion of the Jarzynski's equation was used to calculate the free energy profile along the reaction coordinate and to obtain the potential of mean force (PMF) of the SMD simulations. The three paths chosen presented different energetic barriers related to the disruption of inhibitor interactions with the enzyme. These energy barriers calculated by the simulations are in the same order of magnitude as the bond free energies for similar systems obtained from the experimental inhibition constants.

Keywords: Molecular Dynamics, EPSPs enzyme, Glyphosate, AFM, Potential of Mean Force.

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 Biomoléculas.....	14
1.2 Aminoácidos e Proteínas.....	14
1.3 Enzimas.....	20
1.4 A enzima 5-enolpiruvil-shikimato-3-fosfato sintase (EPSPs).....	22
1.5 Glifosato.....	23
1.6 Microscopia de força atômica (AFM).....	26
2 OBJETIVOS.....	31
2.1 Objetivos Gerais.....	31
2.1.1 Objetivos Específicos.....	31
3 FUNDAMENTAÇÃO TEÓRICA.....	32
3.1 Modelagem Molecular.....	32
3.2 Mecânica Quântica.....	33
3.3 Mecânica Molecular (MM).....	38
3.3.1 Potencial de ligação.....	39
3.3.2 Potencial angular.....	39
3.3.3 Potencial torsional.....	39
3.3.4 Potencial de Lennard-Jones.....	41
3.3.5 Potencial de Coulomb.....	41
3.4 <i>Ensemble</i> Estatístico e Tratamento dos Sistemas Simulados.....	41
3.5 Dinâmica Molecular.....	45
3.6 Métodos de integração da equação de movimento.....	47
3.7 Condições periódicas de contorno.....	50
3.8 Tratamento das interações de curto e longo alcance.....	51
3.9 Modelos de solvatação contínua.....	53
3.10 Dinâmica molecular de sistemas aquosos.....	56
3.11 Acoplamento de temperatura e pressão.....	58
3.12 Minimização de energia.....	60
4 METODOLOGIA.....	62
4.1 Sistemas simulados.....	62
4.2 Parametrização do herbicida glifosato.....	62
4.3 Simulação do sistema EPSPs-GPJ-S3P.....	66
4.4 Cálculo do Potencial Eletrostático e da Energia Livre de Solvatação.....	66

4.5 Dinâmica Molecular Direcional e Cálculo do Potencial de Força Média.....	66
5 RESULTADOS E DISCUSSÕES.....	70
5.1 Caracterização da dinâmica estrutural e do potencial eletrostático da enzima EPSPs em solução aquosa.....	70
5.2 Desenvolvimento de um modelo molecular para o herbicida e parametrização de campos de força.....	74
5.3 Caracterização do sistema EPSPs-S3P-GPJ.....	79
5.4 Descrição atomística e previsões moleculares no processo de detecção do herbicida pelo nanobiossensor.....	84
6 CONCLUSÃO.....	93
7 BIBLIOGRAFIA.....	94

1 Introdução

1.1 Biomoléculas

As biomoléculas possuem características químicas comuns às demais moléculas da natureza. Porém, quando associadas em um sistema biológico, possuem uma dinâmica própria de regulação e síntese, que proporcionam as características de cada ser vivo¹. As biomoléculas, em sua maioria, são compostos orgânicos derivados dos hidrocarbonetos, os quais são compostos formados por um esqueleto de átomos de carbono ligados covalentemente entre si e aos quais estão ligados apenas átomos de hidrogênio, sendo esses esqueletos carbônicos muito estáveis. Os átomos de hidrogênio podem ser substituídos individualmente por uma grande variedade de grupos funcionais para formar famílias diferentes de compostos orgânicos. A “personalidade” química de um composto é determinada pela química de seus grupos funcionais e a disposição destes no espaço tridimensional².

As principais classes de biomoléculas em todos os tipos de células são: proteínas, ácidos nucleicos, polissacarídeos e lipídios, sendo as proteínas as mais versáteis das biomoléculas nos sistemas vivos e servem para funções cruciais em essencialmente todos os processos biológicos. Funcionam como catalisadores, transportam e armazenam outras moléculas, tais como oxigênio, fornecem apoio mecânico e proteção imunitária, geram movimentos, transmitem impulsos nervosos, e controlam o crescimento e a diferenciação³.

1.2 Aminoácidos e Proteínas

As proteínas ocorrem em uma grande variedade, desde peptídeos de tamanho relativamente pequeno até enormes polímeros com pesos moleculares na faixa de milhões, isto em apenas uma célula. As suas subunidades monoméricas relativamente simples fornecem a chave para a estrutura de milhares de proteínas diferentes. Todas as proteínas, sejam das linhagens mais antigas de bactérias, sejam das formas mais complexas de vida, são construídas com o mesmo ubíquo conjunto de 20 aminoácidos (Figura 1), ligados covalentemente em sequências lineares características. Em virtude de cada um desses aminoácidos ter uma cadeia lateral distintiva, a qual determina as suas propriedades químicas, esse grupo de 20 moléculas precursoras pode ser considerado como o alfabeto com o qual a linguagem da estrutura proteica é escrita².

Os aminoácidos são compostos que apresentam, na sua molécula, um grupo amino ($-NH_2$) e um grupo carboxila ($-COOH$); a única exceção é a prolina, que contém um grupo imino ($-NH-$) no lugar do grupo amino. Em pH fisiológico, esses grupos estão na forma

ionizada: -NH_3^+ , -COO^- e -NH_2^+ . Os aminoácidos têm uma fórmula básica comum, na qual os grupos amino e carboxila estão ligados ao carbono α , ao qual também se liga um átomo de hidrogênio e um grupo variável chamado cadeia lateral ou grupo R:

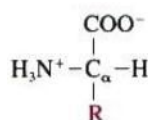
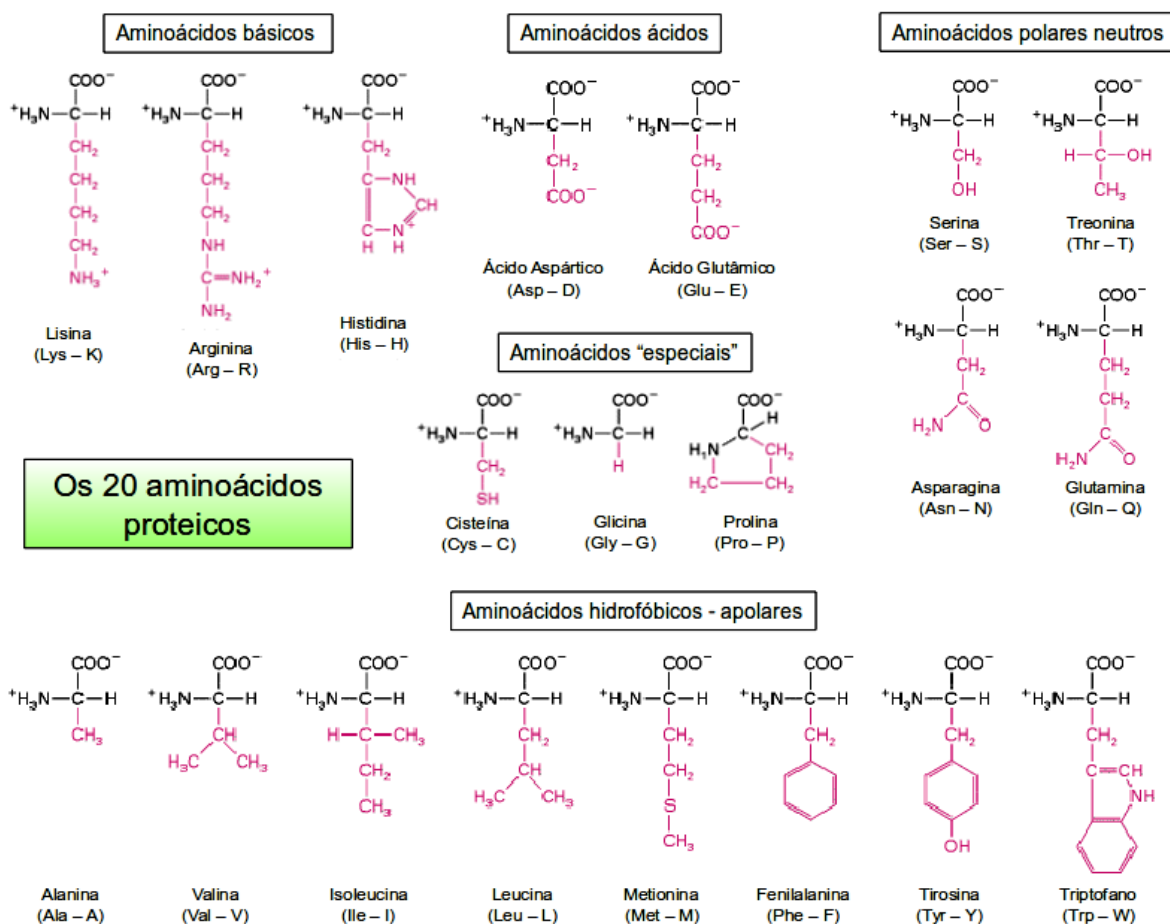


Figura 1 - Estrutura e classificação dos aminoácidos.



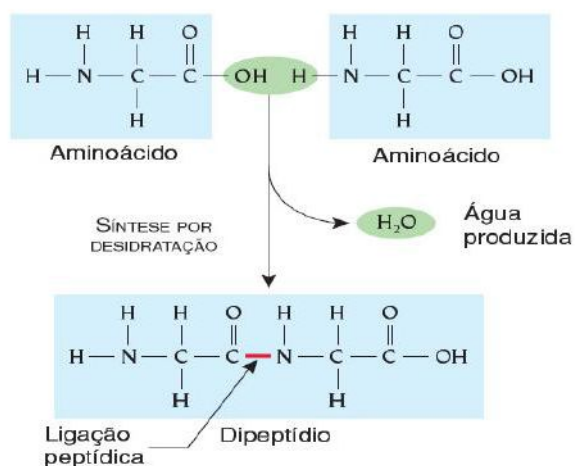
Fonte: <http://slideplayer.com.br/slide/333917/>, acessado em 27/07/2015.

As propriedades das cadeias laterais dos aminoácidos – principalmente o fato de algumas delas terem afinidade pela água e outras não – são importantes para a conformação das proteínas e, portanto, para sua função. Os aminoácidos são classificados, de acordo com a polaridade do grupo R, em duas grandes categorias: aminoácidos apolares (grupo R hidrofóbico) e aminoácidos polares (grupo R hidrofílico)⁴, Figura 1.

Duas moléculas de aminoácidos podem ser unidas covalentemente por meio de uma

ligação amida substituída, chamada de ligação peptídica, para formar um dipeptídeo. Tal ligação é formada por remoção dos elementos da água (desidratação) de um grupo α -carboxila de um aminoácido e de um grupo α -amino de outro (Figura 2)². Uma série de aminoácidos unidos por ligações peptídicas formam uma cadeia peptídica, ou polipeptídica, sendo que esta cadeia é constituída de uma parte que se repete regularmente, chamada de cadeia principal ou arcabouço, e uma parte variável, que corresponde às distintas cadeias laterais (Figura 3). O arcabouço peptídico é rico em potencial formação de ligações de hidrogênio. Cada unidade contém um grupamento carbonila, que é um bom aceptor para estes tipos de ligações e, com exceção da prolina, um grupamento NH, que é um bom doador para este tipo de interação molecular. A maioria das cadeias polipeptídicas naturais contém entre 50 a 2000 aminoácidos, e são comumente referidas como proteínas³.

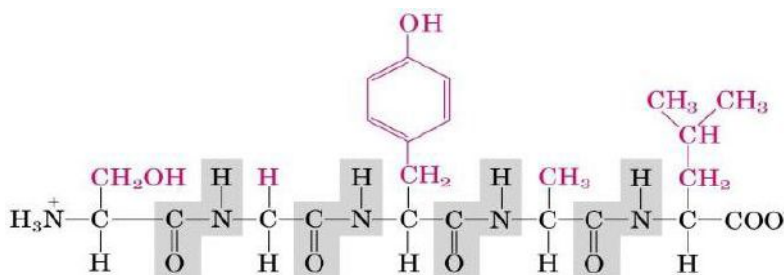
Figura 2 - Ligação peptídica.



Fonte: https://thinkbio.wordpress.com/2011/12/13/materia_organica/, acessado em 28/07/2015.

O fato marcante é que cada proteína tem uma sequência particular de aminoácidos, definida com precisão, sendo esta sequência referida frequentemente como estrutura primária de uma proteína. O exame da geometria da cadeia principal das proteínas revela diversos aspectos importantes.

Figura 3 - Cadeia peptídica: Em cinza a cadeia principal, formada através das ligações peptídicas, e em vermelho as cadeias laterais.



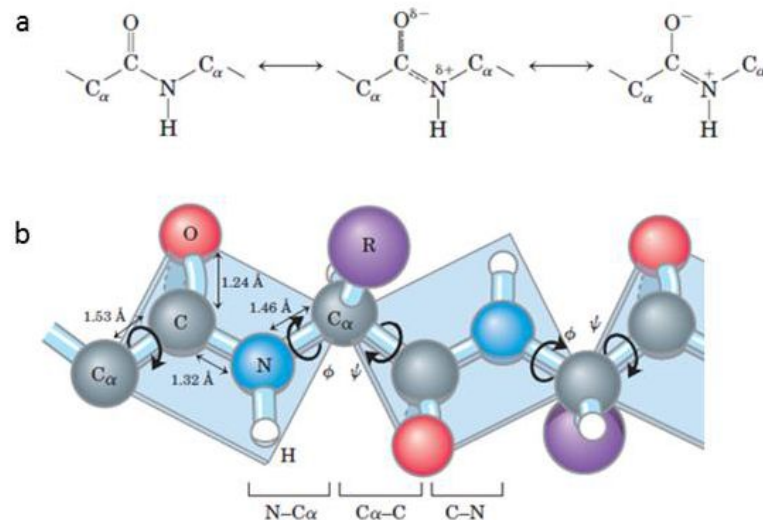
Fonte: <http://www.ebah.com.br/content/ABAAAgEAoAD/proteinas>, acessado em 28/07/2015.

No final da década de 1930, Linus Pauling e Robert Corey realizaram uma série de estudos que levaram aos fundamentos de nossa presente compreensão da estrutura das proteínas. Eles principiaram por uma análise cuidadosa da ligação peptídica. Os carbonos α dos resíduos de aminoácidos adjacentes estão separados por três ligações covalentes, em um arranjo que pode ser representado por $C_\alpha - C - N - C_\alpha$. Estudos de difração de raios X em cristais de aminoácidos, de dipeptídeos e tripeptídeos simples demonstraram que a ligação $C - N$ em um peptídeo é um pouco mais curta que a ligação $C - N$ em uma amina simples e os átomos associados com a ligação peptídica são coplanares. Isso indica uma ressonância ou compartilhamento parcial de dois pares de elétrons entre o oxigênio do grupo carbonila e o nitrogênio do grupo amida (Figura 4-a)².

Desta forma, a ligação peptídica tem um considerável caráter de dupla ligação, constringindo a ligação e automaticamente impedindo a rotação em torno dela, o que explica a sua planaridade (Figura 4-b). As rígidas ligações peptídicas limitam a faixa de conformações que podem ser assumidas por uma cadeia polipeptídica.

Ao contrário das ligações peptídicas, as ligações entre o $N - C_\alpha$ e entre $C_\alpha - C$ são ligações simples puras. As duas unidades peptídicas rígidas adjacentes podem rodar em torno destas ligações, assumindo várias orientações. Esta liberdade de rotação em torno de duas ligações de cada aminoácido permite que as proteínas se enovelem de muitos modos diferentes. Por convenção, os ângulos de ligações resultantes das rotações em C_α são denominados ϕ (fi) para a ligação $N - C_\alpha$ e ψ (psi) para a ligação $C_\alpha - C$.

Figura 4 - Grupo peptídico planar: (a) Caráter de dupla ligação da ligação peptídica; (b) As ligações N—C α e C α —C podem sofrer rotação em ângulos denominados ϕ e ψ , respectivamente.



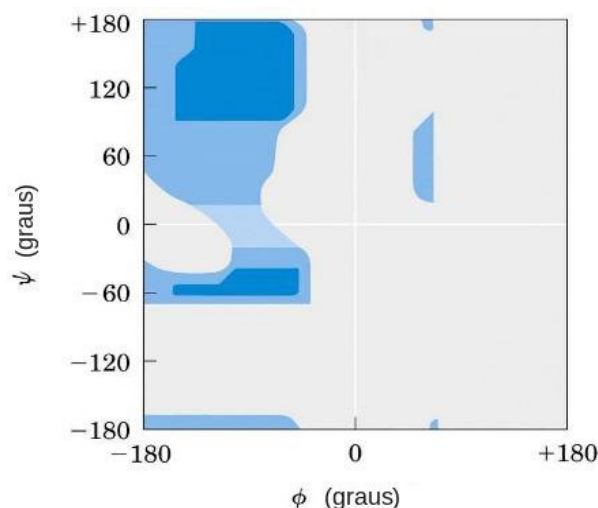
Fonte:²

Novamente por convenção, tanto ϕ como ψ são definidos como sendo iguais a 180° quando o polipeptídeo está em sua conformação totalmente estendida e todos os grupos peptídicos estão no mesmo plano (Figura 4-b). A princípio, ϕ e ψ podem ter qualquer valor entre -180° e $+180^\circ$, mas muitos valores são inatingíveis por interferência estérica entre átomos na estrutura polipeptídica e cadeias laterais de aminoácidos^{2,3}.

Os valores das combinações estruturais permitidas podem ser visualizados em um gráfico com duas dimensões, chamado de diagrama de Ramachandran (Figura 5)⁵.

As proteínas apresentam arranjos tridimensionais com dobramentos regulares, sendo que estas estruturas são estabilizadas por ligações de hidrogênio entre o oxigênio carbonil de uma ligação peptídica e o hidrogênio amida de uma outra ligação peptídica próxima ($-\text{NH}\cdots\text{O}=\text{C}-$). Duas organizações são particularmente estáveis: o enrolamento da cadeia ao redor de um eixo e a interação lateral de segmentos de uma cadeia polipeptídica ou de cadeias diferentes. Estas conformações são denominadas, respectivamente, de α -hélice e folha β pregueada (Figura 6)^{4,6}.

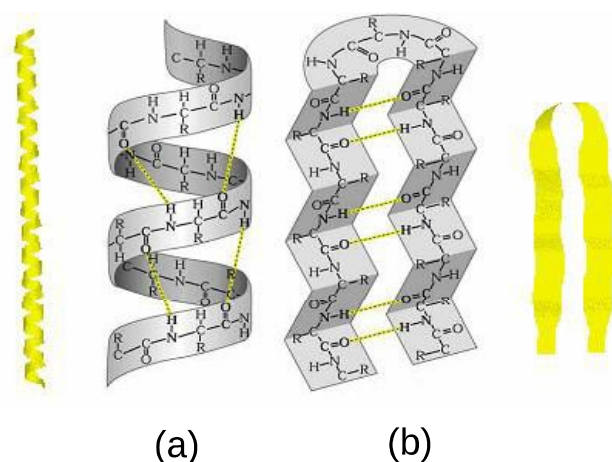
Figura 5 - Diagrama de Ramachandran para resíduos de L-Ala. As conformações dos peptídeos são definidas pelos valores de ψ e ϕ . As áreas em azul-escuro refletem conformações que não envolvem sobreposições estéricas, permitidas; o azul-médio indica conformações permitidas nos limites extremos para os contatos atômicos desfavoráveis; a área azul-clara reflete conformações que são permitidas se houver uma pequena flexibilidade nos ângulos de ligação e área em cinza conformações proibidas.



Fonte:²

A capacidade de polímeros biológicos como as proteínas se enovelarem em estruturas bem definidas é termodinamicamente notável. No contexto da estrutura proteica, o termo estabilidade pode ser definido como sendo a tendência à manutenção de uma conformação nativa. As proteínas nativas são apenas ligeiramente estáveis; a variação de energia livre, ΔG , que separa os estados enovelado e desenovelado em proteínas típicas sob condições fisiológicas se situa na faixa de apenas 20 a 65 KJ/mol. A entropia favorável associada ao grande número de conformações na forma desdobrada se opõe ao enovelamento e tem de ser vencida por interações favorecendo a forma enovelada. Assim, polímeros muito flexíveis com um grande número possível de conformações não se enovelam em estruturas únicas. A rigidez da unidade peptídica e o conjunto restrito de ângulos ϕ e ψ limita o número de estruturas acessíveis à forma não enovelada, o suficiente para permitir que ocorra o enovelamento da proteína^{2,3}. A estrutura nativa, e portanto, a estabilidade de uma proteína é fundamental para a sua atividade biológica.

Figura 6 - Estruturas em (a) α -hélice e (b) folha β pregueada.



Fonte: <http://gracieteoliveira.pbworks.com/w/page/50091415/A%20c%C3%A9lula>, acessado em 03/08/2015.

As proteínas são, sem dúvida nenhuma, as biomoléculas de maior número de funções biológicas em um organismo vivo, dentre elas, a indispensável função catalisadora exercida pelas enzimas, sem a qual não seria possível a maioria das reações celulares.

1.3 Enzimas

Em seres vivos, a maioria das reações químicas são catalisadas por enzimas, que são proteínas com funções catalíticas que determinam o perfil de transformações químicas. As características mais impressionantes das enzimas são a eficiência catalítica extraordinária e o alto grau de especificidade. As enzimas são altamente específicas tanto na reação catalisada como na sua escolha de reagentes, os quais são chamados de substratos. Além disso, a atividade catalítica de muitas enzimas depende da presença de pequenas moléculas denominadas co-fatores ou coenzimas, embora seu papel exato varie com o co-fator e com a enzima. Os co-fatores ou coenzimas são constituídos por metais ou por moléculas orgânicas pequenas.

O centro ativo de uma enzima é a região que se liga aos substratos (e ao co-fator, se houver algum) e que contém os radicais de aminoácidos que participam diretamente na geração e quebra de ligações. Estes radicais são chamados de grupamentos catalíticos. Em essência, a interação da enzima com o substrato no centro ativo promove a formação do estado de transição³.

Os aumentos de velocidade de reação que variam entre 5 a 17 ordens de magnitude e a discriminação de substratos com estruturas similares promovidos pelas enzimas, podem ser

relacionados com dois princípios distintos, mas interligados.

O primeiro diz respeito aos rearranjos das ligações covalentes durante a reação enzimática. Reações químicas dos mais variados tipos ocorrem entre o substrato e os grupos funcionais da enzima (cadeias laterais de aminoácidos específicos, co-fatores metálicos e coenzimas). Os grupos funcionais catalíticos das enzimas podem formar uma ligação covalente transitória com um substrato e ativá-lo para a reação ou, então, algum grupo pode ser transferido transitoriamente do substrato para um grupo da enzima. Em muitos casos, essas reações somente ocorrem no centro ativo da enzima. Elas diminuem a energia de ativação (e, portanto, aceleram a reação), propiciando um caminho reacional alternativo de energia mais baixa.

As interações não-covalentes entre a enzima e o substrato, são responsáveis pelo segundo princípio. As interações fracas, não-covalentes, entre o substrato e a enzima são responsáveis pela maior parte da energia necessária para diminuir a energia de ativação. O fator, que de fato distingue as enzimas da maioria dos catalisadores não-enzimáticos é a formação de um complexo enzima-substrato (ES) específico. A interação entre o substrato e a enzima nesse complexo é mediada pelas mesmas forças que estabilizam a estrutura proteica, incluindo as ligações de hidrogênio, interações dipolo-dipolo e forças de van der Waals. A formação de cada interação fraca no complexo ES é acompanhada por uma pequena liberação de energia livre que garante o grau de estabilidade para a interação. A energia derivada da interação ES é chamada de energia de ligação. O seu significado vai além de uma simples estabilização da interação enzima-substrato. A energia de ligação é a maior fonte de energia livre usada pelas enzimas para diminuir a energia de ativação das reações, ou seja, as interações entre a enzima e o substrato fornecem a maior parte da força que dirige a catálise enzimática, contribuindo tanto para a especificidade quanto para a catálise².

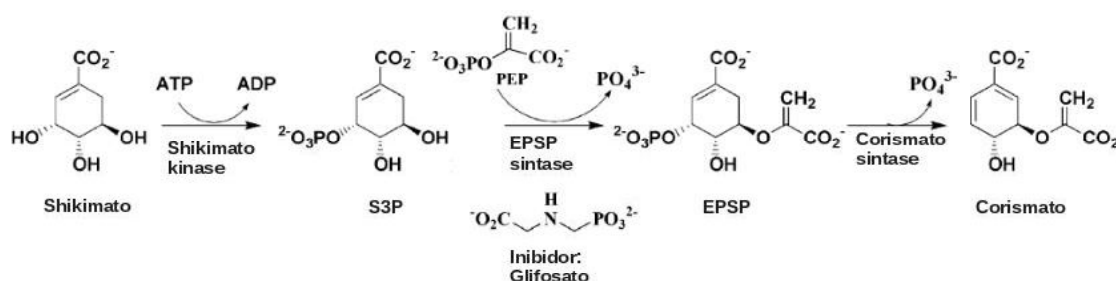
A atividade enzimática pode ser interrompida ou inibida por um grande número de substâncias, genericamente chamadas de inibidores. Como a ação enzimática efetivamente controla o metabolismo, a possibilidade de inibir reações enzimáticas é um campo aberto para diversas aplicações. Muitos medicamentos de uso corrente na prática terapêutica baseiam suas propriedades na inibição específica de certas enzimas. Muitos inseticidas e herbicidas têm nos inibidores enzimáticos o seu princípio ativo⁴.

O presente trabalho está relacionado diretamente com a enzima catalisadora 5-enolpiruvil-shikimato-3-fosfato sintase (EPSPs) e o seu inibidor específico (N-(fosfonometil) glicina, conhecido com glifosato, e serão discutidos mais detalhadamente a seguir.

1.4 A enzima 5-enolpiruvil-shikimato-3-fosfato sintase (EPSPs)

A rota shikimato constitui uma parte do metabolismo encontrado somente em micro-organismos e plantas, nunca em animais. Esta rota é essencial para a biossíntese do corismato, o precursor para formação dos aminoácidos aromáticos fenilalanina, tirosina e triptofano. A enzima shikimato quinase catalisa a sexta etapa da rota convertendo shikimato a shikimato-3-fosfato (S3P). A parte enolpiruvil do fosfoenol-piruvato (PEP) é transferido para o grupo 5-hidroxila do S3P, produzindo o 5-enolpiruvil-shikimato-3-fosfato (EPSP). Esta reação é catalisada pela enzima 5-enolpiruvil-shikimato-3-fosfato sintase (EPSPs). A etapa final do caminho shikimato é a conversão da EPSP a corismato pela corismato sintase (Figura 7)^{7,8}.

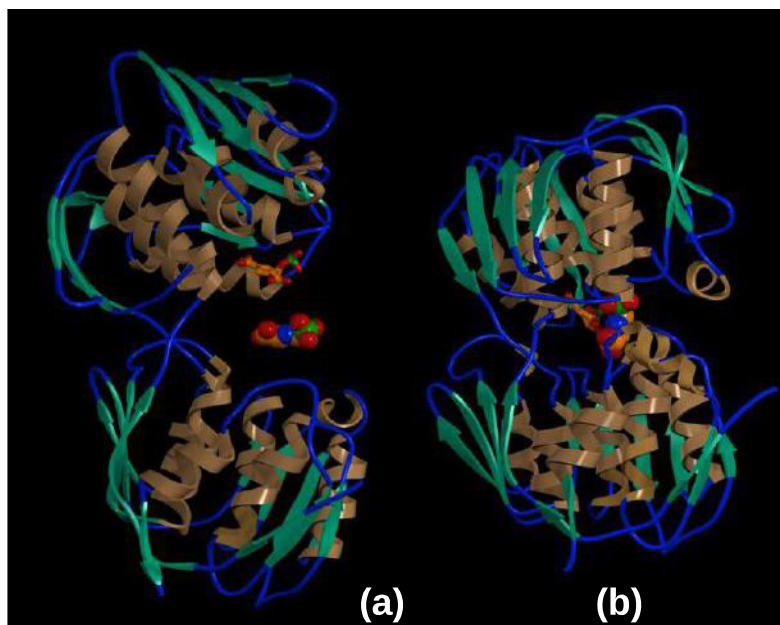
Figura 7 - Conversão do shikimato a corismato na rota shikimato.



Fonte: adaptado de ⁸

A estrutura terciária de uma única molécula de EPSP sintase da bactéria *Escherichia coli* (*E. coli*) está ilustrada na Figura 8. A estrutura consiste de dois domínios globulares distintos que são aproximadamente hemisféricos, cada um com um raio de cerca de 25 Å, possuindo um total de 427 resíduos de aminoácidos. Distribuído sobre o eixo de cada hemisfério estão três hélices paralelas de três a quatro voltas cada; as superfícies de domínio incluem três folhas β , a maioria com quatro fitas, e três hélices paralelas que possuem faces acessíveis ao solvente. Estes constituem os principais segmentos da estrutura secundária da enzima EPSP sintase^{9,10}. A enzima EPSPs é o único alvo do inibidor glifosato, o ingrediente ativo do herbicida Roundup® da Monsanto. A reação da EPSPs com o glifosato, na presença do substrato S3P é alvo de diversos estudos há vários anos.

Figura 8 - Representação (em cartoon) da conformação da enzima EPSPs; (a) Estado não ligado (aberto) e (b) Estado ligado (fechado). O S3P e o glifosato são representados nos modelos (bola e bastão) e (van der Waals-VDW), respectivamente.



Fonte: <http://www.artn.com/Monsanto>, acessado em 03/08/2015.

1.5 Glifosato

O elevado crescimento da atividade agrícola está associado à incorporação de várias tecnologias ao processo de modernização da agricultura, dentre elas, destaca-se o uso de defensivos agrícolas sendo o Brasil o maior consumidor de agrotóxicos do mundo¹¹. As variáveis que constroem esse indicador envolvem a área plantada em hectares(ha) e a quantidade de agrotóxicos consumidos de acordo com a classe, discriminados segundo as principais classes de uso (herbicidas, fungicidas, inseticidas, bactericidas, dentre outros)¹².

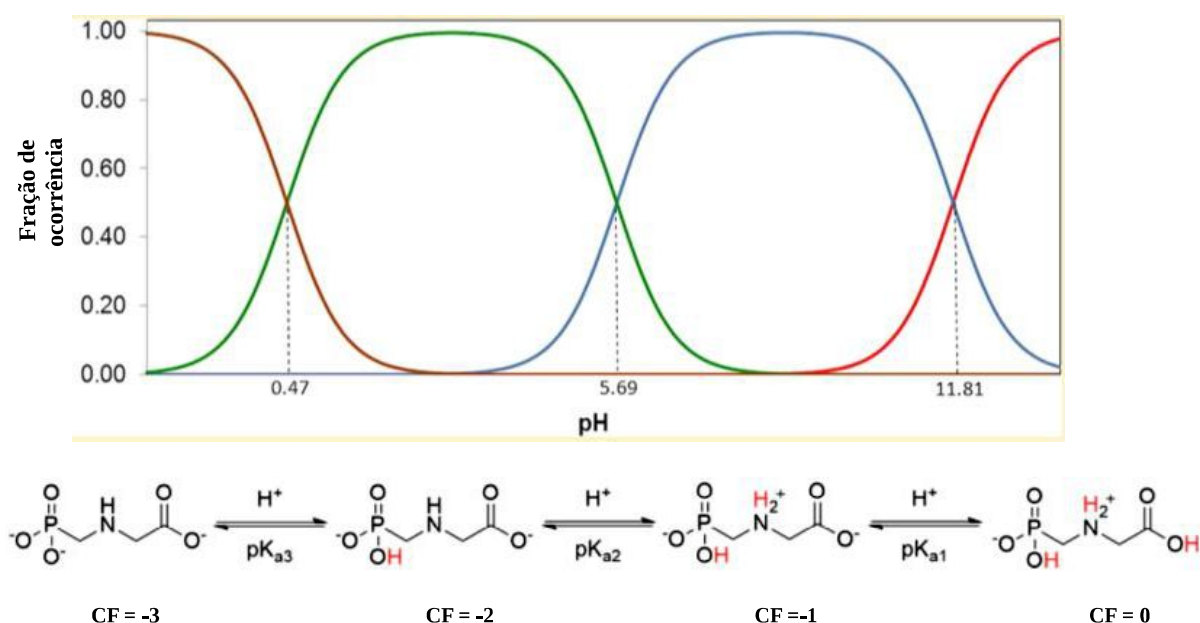
Dentre os herbicidas, o glifosato, (N-(fosfonometil) glicina é o ingrediente ativo mais utilizado em todo o mundo. A propriedade herbicida desta molécula foi descoberta pela Monsanto em 1970 e a primeira formulação comercial foi lançada nos Estados Unidos em 1974, com o nome comercial de Roundup®. É utilizada para o controle de plantas daninhas nas áreas agrícolas, industriais, florestais, residenciais e ambientes aquáticos, de acordo com os registros obtidos em cada país. O glifosato é um herbicida pós-emergente, pertencente ao grupo químico das glicinas substituídas, classificado como não-seletivo e de ação sistêmica. Apresenta largo espectro de ação, o que possibilita um excelente controle de plantas daninhas anuais ou perenes, tanto de folhas largas como estreitas em culturas de arroz, cana-de-açúcar, café, milho, soja, dentre outras¹³.

O tecido vivo das plantas absorve o herbicida e o transloca, via floema, através da planta para raízes e rizomas, e sua ação inibe a enzima específica 5-enolpiruvil-shikimato-3-fosfato sintase (EPSPs) (Figura 8). Após o tratamento com o glifosato as plantas morrem lentamente, em dias ou semanas, e nenhuma parte dela sobrevive devido ao transporte por todo o sistema¹⁴.

Numerosos estudos demonstraram que o glifosato forma preferencialmente um complexo ternário estável com a enzima EPSPs e o substrato S3P. Este complexo ternário representa provavelmente a forma real da ligação do glifosato na enzima, que é responsável pela sua atividade herbicida nas ervas daninhas¹⁵⁻¹⁸.

Em condições ambientais, tanto glifosato quanto seus sais são sólidos cristalinos, muito solúveis em água (12 g/L a 25 °C, para glifosato) e quase insolúveis em solventes orgânicos comuns, tais como acetona e etanol¹⁹. Os valores de pK_a encontrados para o glifosato são: $pK_1 = 0,47$; $pK_2 = 5,69$; $pK_3 = 11,81$. Tais constantes de dissociação indicam o grau de dissociação do herbicida em função do $pH^{20,21}$ (Figura 9).

Figura 9 - Fração de ocorrência das espécies de glifosato construído a partir dos valores de pK_a (marrom, composto com carga formal igual a 0; verde, composto com carga formal igual a -1; azul, composto com carga formal igual a -2; vermelho composto com carga formal igual a -3 ; e representação dos equilíbrios.

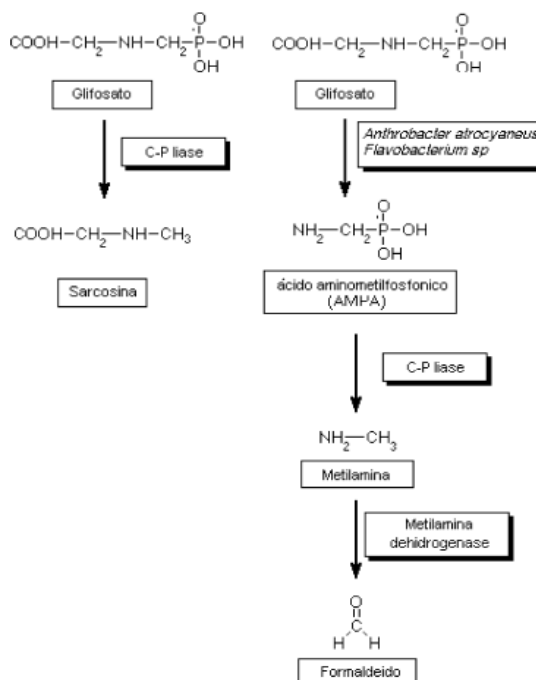


Fonte: Adaptado de ^{20,21}.

O glifosato possui duas principais rotas de degradação microbiológica (Figura 10). A primeira consiste na clivagem oxidativa da ligação carbono nitrogênio produzindo o ácido aminometilfosfônico (AMPA), o qual pode ser degradado a fosfato inorgânico. A segunda rota consiste na clivagem da ligação fósforo-carbono por ação da bactéria *Agrobacterium radiobacter* ou da *Enterobacter aeroneges* (enzima C-P liase) produzindo a sarcosina²².

As concentrações mais altas de glifosato e seu metabólito, o AMPA, tem sido encontradas em folhagens novas. A aplicação de glifosato pode resultar na presença de resíduos tanto na colheita quanto em animais usados na alimentação humana. No ambiente, as concentrações mais altas de ambos os compostos foram encontradas no solo^{19,23}.

Figura 10 - Rotas de decomposição microbiológica do glifosato.



Fonte:¹⁹.

A toxicidade aguda do glifosato é considerada baixa, pertencente à classe IV (pouco tóxico), segundo classificação da Agência Nacional de Vigilância Sanitária (ANVISA)²⁴. Embora a toxicidade aguda seja considerada baixa, diversos estudos sobre os efeitos epidemiológicos, em humanos e outros animais, deste herbicida, são constantemente realizados. Estudos sobre a indução à apoptose e necrose em células humanas²⁵, efeitos do glifosato em embriões de anfíbios²⁶, efeitos na atividade da colinesterase em molusco²⁷, alterações nos níveis de citocina em carpas²⁸, alteração de células relacionadas à doença de Parkinson²⁹, além de estudos de citotoxicidade e genotoxicidade em células humanas³⁰. A

avaliação do efeito carcinogênico recebe atenção especial, uma vez que gera bastante controvérsia, estudos relatam que não há evidências de um padrão consistente que seria indicativo de uma relação causal entre qualquer tipo de câncer e exposição ao glifosato³¹, porém, há autores que relatam que o herbicida induz o desenvolvimento de tumor em peles de ratos³², além de estudos que demonstram que a exposição ao glifosato pode representar um risco de câncer de mama³³. No entanto, a ANVISA anunciou oficialmente que reavaliará a liberação do uso do glifosato no Brasil. A postura foi motivada pelo artigo publicado na Revista “The Lancet” em março de 2015 pela Agência Internacional de Pesquisa sobre o Câncer (Iarc), órgão da Organização Mundial da Saúde (OMS), classificando a carcinogenicidade de cinco agrotóxicos dentre eles o glifosato³⁴. Os autores classificaram o glifosato como provavelmente cancerígeno para os seres humanos³⁵. A investigação de pesticidas depara-se com problemas analíticos complexos, ocasionados pela gama de substâncias existentes a serem determinadas em distintos produtos agrícolas, possivelmente presentes em níveis traços³⁶. Diversas técnicas analíticas têm sido utilizadas na detecção e quantificação de pesticidas. Métodos baseados em cromatografia líquida³⁷, cromatografia gasosa³⁸, métodos analíticos óticos³⁹, técnicas baseadas em voltametria pulsada⁴⁰, dentre outros. Os biossensores surgiram como métodos analíticos alternativos. São arranjos que apresentam, em geral, elevadas sensibilidade e seletividade, com tempo de resposta bastante curto e mínimo tratamento da amostra.

O biossensor é uma ferramenta analítica que pode ser definida como a combinação de um elemento de reconhecimento biológico (como por exemplo, enzima, célula, organela, anticorpo, entre outros) e um transdutor físico, em íntimo contato entre eles. O papel do transdutor é relacionar a concentração de um determinado analito a uma propriedade química ou física, que será sequencialmente detectada, convertida em um sinal elétrico/óptico, que é amplificado⁴¹.

Uma importância particular são os nanobiossensores para detectar herbicidas, especialmente os inibidores enzimáticos, obtidos pela deposição de uma camada receptora (proteína) em microcantileveres, com analitos detectados a uma concentração tão baixa quanto 10^{-18} mol.L⁻¹ usando microscopia de força atômica (AFM, do inglês, Atomic Force Microscopy)⁴².

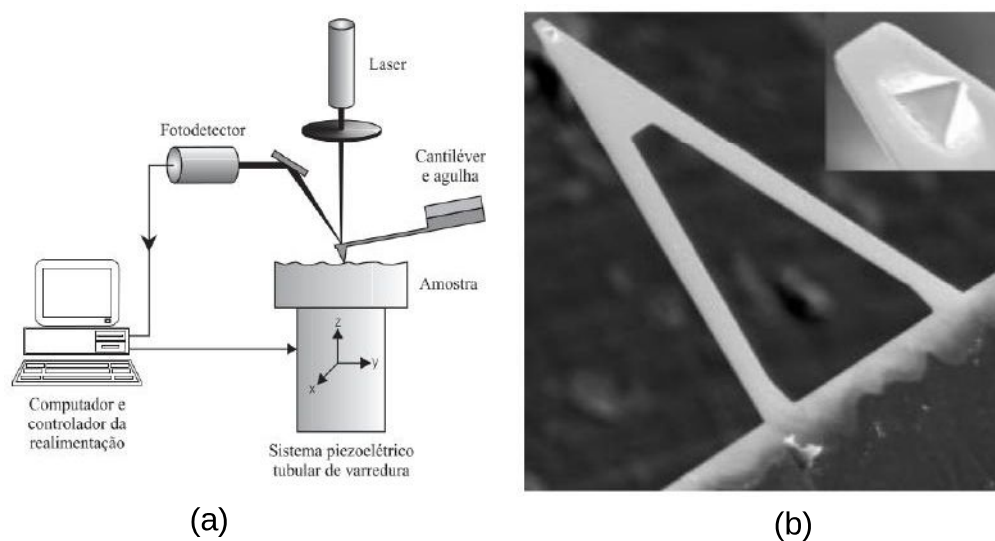
1.6 Microscopia de força atômica (AFM)

O microscópio de força atômica (AFM) foi inventado em 1986 por Binnig, Quate e Berger⁴³. O principal ímpeto para o AFM foi medir pequenas forças. Eles previram, que com a

melhoria da instrumentação, o equipamento seria capaz de identificar forças tão pequenas quanto 10^{-18} N. Com esse limite de detecção, ele poderia detectar até mesmo forças intermoleculares fracas.

O princípio de funcionamento do AFM baseia-se na varredura da superfície da amostra por uma ponta piramidal (ponteira) de alguns micrômetros de comprimento (100 a 200 μm) e geralmente com menos de vinte nanômetros de diâmetro, integrada em um cantiléver flexível. A sonda (ponteira + cantiléver) é sempre o componente básico do AFM e, para alcançar resolução atômica, a ponta tem que terminar em um conjunto de átomos. A força entre a ponta e a superfície da amostra faz com que o cantiléver se aproxime ou se afaste e essa deflexão é proporcional à força de interação⁴⁴ (Figura 11).

Figura 11 - (a) Ilustração do princípio de funcionamento do AFM; (b) Microscopia eletrônica de varredura de um cantiléver de nitreto de silício com agulha acoplada (em detalhe).



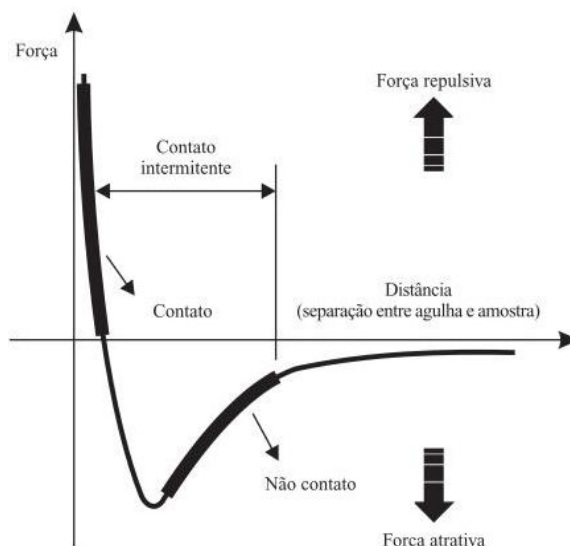
Fonte: ⁴⁵

Esta varredura é feita, por intermédio de um sistema piezoelétrico, com deslocamento, nas posições x , y e z com precisão de décimo de Angstrom, o que se dá através da variação da tensão aplicada no mesmo. O deslocamento é controlado por um circuito de realimentação cuja função é manter a força e/ou altura constante. Um sistema ótico, com feixe a laser e um fotodetector, determinam o quanto o cantiléver deflete devido à topografia da amostra. Com os dados da deflexão do cantiléver nos eixos x , y e z , reconstrói-se a imagem por intermédio de um software adequado⁴⁵.

A força mais comumente associada à microscopia de força atômica é a força de van der Waals. A dependência dessa força sobre a distância da sonda e a amostra descreve duas regiões de interação (Figura 12).

Na zona conhecida como região ou modo de contato, o cantiléver é mantido a uma distância menor que poucos angstroms da superfície da amostra e a força interatômica entre o cantiléver e a amostra é repulsiva. Neste modo repulsivo, a sonda faz um leve “contato físico” com a amostra produzindo imagens com alta resolução, mas a compressão e as forças geradas, entre a ponta e a superfície, podem causar danos à amostra, o que é especialmente prejudicial às amostras biológicas que são sensíveis e nem sempre fortemente aderidas ao substrato.

Figura 12 - Curva de força pela distância de separação entre a agulha e amostra.



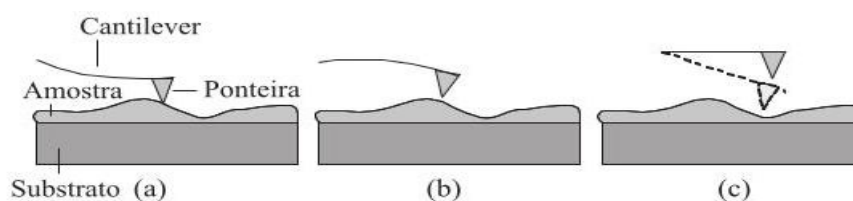
Fonte:⁴⁵.

No modo de não-contato, o cantiléver é mantido na ordem de 10-100 angstroms da superfície da amostra e a força interatômica entre o cantiléver e a amostra é atrativa (em grande parte como resultado das interações de longo alcance de van der Waals). Neste caso, a ponta oscila em alta frequência (100 kHz a 1 MHz), a poucos nanômetros acima da superfície e a força total entre o cantiléver e a amostra nesse regime é muito pequena, geralmente em torno de 10^{-12} N. Este regime de não-contato não sofre os efeitos de degradação da sonda ou da amostra que muitas vezes são observadas no modo contato, no entanto se algumas monocamadas de água condensada estiverem presentes na superfície da amostra rígida, por um instante, as imagens podem se divergir bastante. Na operação no modo contato a sonda penetrará na camada líquida e gerará a imagem da superfície subjacente, enquanto no modo

não-contato o AFM refletirá a imagem da superfície da camada líquida.

Estas limitações têm sido contornadas com a utilização do modo intermitente. Este modo é similar ao não-contato, exceto pelo fato de que a ponta vibrante fica mais próxima da amostra, de forma que tenha um contato intermitente e é utilizado para contornar as limitações impostas pelo modo contato. A comparação das imagens nos modos contato e intermitente mostra que as superfícies são menos modificadas no modo intermitente^{44,46}. Os modos de operação em microscopia de força atômica são ilustrados na Figura 13.

Figura 13 - Representação esquemática dos modos de operação em AFM: (a) modo contato, (b) modo não-contato e (c) modo intermitente.



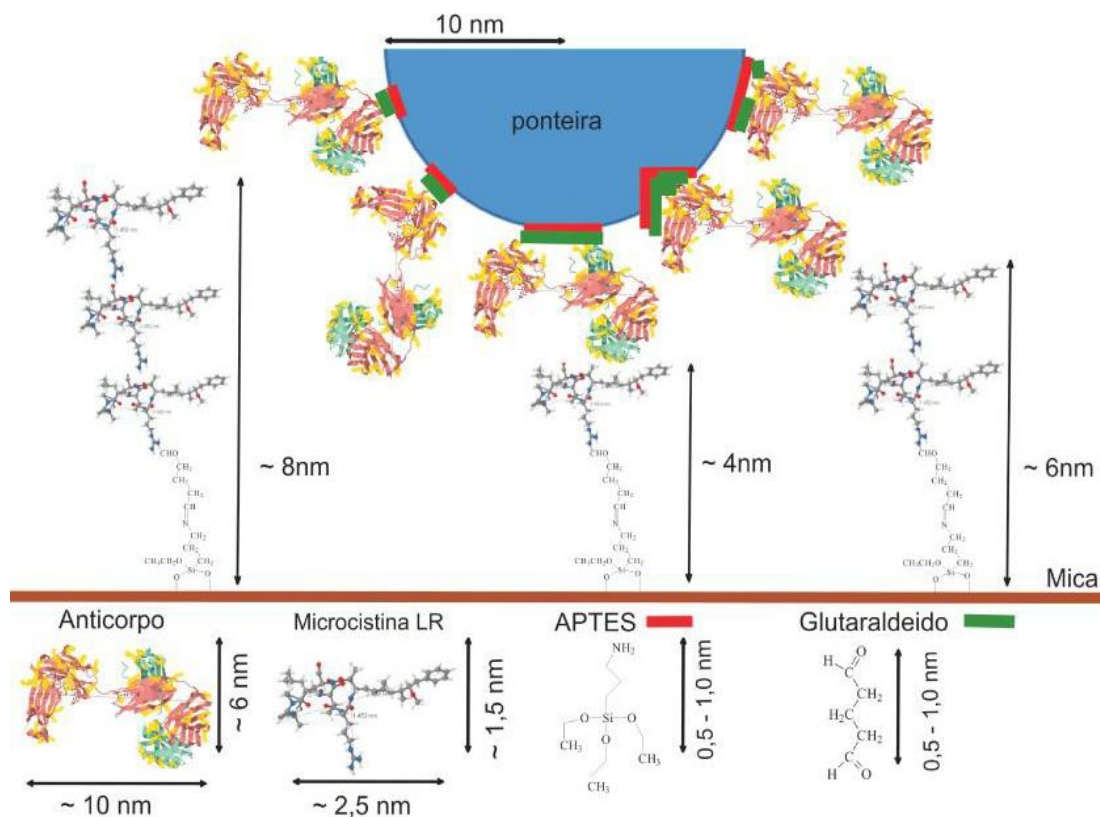
Fonte: ⁴⁴

A microscopia de força atômica é uma poderosa ferramenta utilizada no estudo da morfologia de polímeros⁴⁷, em sistemas biológicos auxiliando na compreensão dos fenômenos biofísicos/químicos envolvidos⁴⁸⁻⁵¹, além de explorar interações de reconhecimento molecular em eventos biológicos ao nível de interações intermoleculares específicas, atuando como nanossensores⁵²⁻⁵⁵, dentre outras diversas aplicações.

Quando o cantiléver é empregado como sensor de força em AFM, podem ser usadas curvas de distância para pesquisar as forças que contribuem para deflexão do cantiléver (ou seja, forças de van der Waals e eletrostáticas) e uma variedade de forças de interações biológicas e químicas não-covalentes podem ser mapeadas diretamente em função da distância ponta-amostra, na qual é baseada na capacidade de medida de força do AFM. Geralmente a ponteira da sonda é funcionalizada quimicamente com o objetivo de promover uma ligação covalente para imobilizar ou adsorver fortemente uma molécula sensorial (enzima ou anticorpo na superfície da ponta). Na maioria das vezes, as interações específicas entre esse sistema e o substrato (amostra) são reversíveis, permitindo o monitoramento contínuo de amostras⁵⁶ (Figura 14).

A força de interação é calculada pela lei de Hooke: $F=k\Delta z$, onde F é força de interação, Δz é deflexão do cantiléver provocada pela interação e k , a constante elástica do cantiléver, que depende de sua geometria e do material de construção⁴⁴.

Figura 14 - Esquema representativo para o sistema ponta funcionalizada-substrato em um AFM.



Fonte: ⁵⁷.

A utilização de ponteiros de AFM quimicamente funcionalizados, a qual também é chamada de microscopia de força química (CFM, do inglês Chemical Force Microscopy), intensifica uma interação intermolecular específica e suprime outras interações interferentes, analisando seletivamente a superfície de grupos funcionais de interesse em micro ou nanoescala, apresentando limite de detecção baixo, em nível molecular⁵⁸.

O ponto negativo em relação a essa técnica, está na necessidade de se conhecer detalhadamente o sistema de interesse, uma vez que se deve escolher as moléculas apropriadas para o reconhecimento molecular, além de avaliar qual região será necessária ficar mais exposta para interagir especificamente com o analito e desta forma, escolher a melhor forma de imobilizar ou funcionalizar essas moléculas na ponta do cantiléver. A grande vantagem de se ter uma análise em escala nanométrica também cria dificuldades no controle da técnica.

Com isso, diversos experimentos precisam ser desenvolvidos para alcançar as condições ideais, o que tem sido uma motivação para o uso de técnicas de modelagem molecular para prever as características do sistema específico, uma vez que, podem elucidar os princípios subjacentes das interações em vários níveis de detalhe.

2 Objetivos

2.1 Objetivos Gerais

A finalidade do estudo é oferecer um aporte teórico suficiente, a nível atômico, para uma possível aplicação no desenvolvimento de um nanobiossensor para a detecção do herbicida glifosato, utilizando um microscópio de força atômica (AFM).

2.1.1 Objetivos Específicos

- Emular um sistema inibidor-enzima, mais especificamente, herbicida glifosato (GPJ) e a enzima 5-Enolpiruvil-Shikimato-3-Fosfato Sintase (EPSPs) via modelagem molecular.
- Analisar a conformação estrutural em solução aquosa e consequentemente a estabilidade da enzima, além da distribuição de cargas na superfície da EPSPs.
- Parametrizar os parâmetros dos substratos para a adição ao campo de força utilizado nas simulações de dinâmica molecular.
- Com o propósito de emular o experimento de detecção do herbicida por microscopia de força atômica, simulações de dinâmica molecular direcional (SMD) serão realizadas para identificar e quantificar as interações entre o GPJ e o sítio ativo da enzima.
- Calcular o Potencial de Força Média (PMF) a partir das simulações de SMD para obtenção da energia livre absoluta de interação entre o herbicida GPJ e a enzima EPSPs.

3 Fundamentação Teórica

3.1 Modelagem Molecular

Químicos do século XIX estavam cientes da conectividade e da geometria básica das moléculas e portanto, das suas fórmulas estruturais, mas não eram capazes de quantificar as estruturas das moléculas em uma base métrica. Além de ligações químicas, estavam cientes das interações de van der Waals, interações eletrostáticas, impedimentos estéricos, conjugação de Kekulé e interações doador-receptor. No entanto, informações detalhadas da estrutura eletrônica e molecular eram desconhecidas.

O século XX trouxe dois avanços importantes. Com o desenvolvimento das técnicas de difração, o arranjo dos átomos pôde ser determinado com uma base métrica. Dependendo da dimensão da molécula, a qualidade do cristal e da precisão do experimento, comprimentos de ligação bem definidos, ângulos de valência e torcionais, bem como contatos não-ligados, puderam ser determinados. O outro desenvolvimento com consequências de longo alcance foi a equação de Schrödinger.

Desta forma, surge a modelagem molecular como uma ferramenta que fornece uma maneira de explorar sistematicamente os padrões estruturais, dinâmicos e termodinâmicos, testar e desenvolver hipóteses, interpretar e ampliar os dados experimentais, ajudar a compreender melhor e expandir as leis básicas que governam a flexibilidade, a função e a estrutura molecular. Por conseguinte, se for possível desenvolver algoritmos que sejam capazes de calcular uma estrutura com uma determinada estequiometria e conectividade, é possível desenvolver algoritmos para o cálculo das propriedades moleculares baseados na estrutura calculada e vice-versa. Porém, antes que qualquer estudo computacional sobre as propriedades moleculares possa ser realizado, um modelo molecular precisa ser estabelecido^{59,60}.

As duas aproximações que têm sido usadas predominantemente em estudos de modelagem molecular são a aproximação clássica, que inclui os métodos da mecânica molecular e da dinâmica molecular, e a aproximação quântica, que inclui os métodos *ab initio* e semi-empíricos. De modo geral, a escolha entre estas aproximações depende das propriedades que se deseja avaliar, da precisão desejada e da capacidade computacional disponível para a realização dos cálculos⁶¹.

Os postulados fundamentais da mecânica quântica afirmam que os sistemas microscópicos são descritos por funções de onda que caracterizam completamente todas as propriedades físicas do sistema. Em particular, existem operadores mecânico-quânticos correspondentes para cada observação física que, quando aplicados à função de onda, permite

prever a probabilidade de um sistema exibir um valor ou intervalo de valores (escalar, vetorial, etc) para aquela observação⁶².

3.2 Mecânica Quântica

A mecânica quântica descreve as moléculas em termos de interações entre o núcleo e os elétrons, e a geometria molecular em termos de arranjos de mínimos de energia dos núcleos. Todos os métodos mecânico-quânticos fundamentalmente remontam da equação de Schrödinger, na qual para o caso especial do átomo de hidrogênio (partícula única em três dimensões) pode ser resolvido de maneira exata (equação 1),

$$\left[-\frac{1}{2} \nabla^2 - \frac{Z}{r} \right] \psi(r) = E \psi(r) \quad (1)$$

onde, a quantidade entre chaves representa a energia potencial e cinética de um elétron a uma distância r do núcleo de carga Z (1 para o hidrogênio). E é a energia eletrônica em unidades atômicas e Ψ , uma função das coordenadas dos elétrons r , é uma função de onda descrevendo o movimento do elétron.

A equação de Schrödinger generalizada para sistemas complexos é representada abaixo:

$$H\Psi = E\Psi \quad (2)$$

Aqui Ψ , é uma função de onda para vários elétrons e H é chamado operador Hamiltoniano, na qual as unidades atômicas são dadas por:

$$\hat{H} = -\frac{1}{2} \sum_i^{\text{elétrons}} \nabla_i^2 - \frac{1}{2} \sum_A^{\text{núcleo}} \frac{1}{M_A} \nabla_A^2 - \sum_i^{\text{elétrons}} \sum_A^{\text{núcleo}} \frac{Z_A}{r_{iA}} + \sum_{i < j}^{\text{elétrons}} \frac{1}{r_{ij}} + \sum_{A < B}^{\text{núcleo}} \frac{Z_A Z_B}{R_{AB}} \quad (3)$$

onde Z é a carga nuclear, M_A é a razão entre a massa do núcleo A e a massa do elétron, R_{AB} é a distância entre o núcleo A e o núcleo B , r_{ij} é a distância entre os elétrons i e j e r_{iA} é a distância entre o elétron i e o núcleo A .

A equação de Schrödinger para sistemas complexos não pode ser resolvida exatamente (ou, pelo menos, ainda não foi resolvida), mesmo para um sistema simples de dois elétrons, tais como o hélio ou a molécula de hidrogênio.

Os problemas encontrados para resolver a equação de Schrödinger para sistemas

complexos resultaram no desenvolvimento de várias aproximações para fornecer métodos práticos. Uma maneira de simplificar a equação para sistemas moleculares é assumir que o núcleo se move muito lentamente quando comparado com a velocidade na qual os elétrons se movem (velocidade da luz). Esta é chamada aproximação de Born-Oppenheimer e conduz à equação “eletrônica” de Schrödinger.

$$\hat{H}^{el} \Psi^{el} = E^{el} \Psi^{el} \quad (4)$$

$$\hat{H}^{el} = -\frac{1}{2} \sum_i^{elétrons} \nabla_i^2 - \sum_i^{elétrons} \sum_A^{núcleo} \frac{Z_A}{r_{iA}} + \sum_{i < j}^{elétrons} \frac{1}{r_{ij}} \quad (5)$$

O termo na equação (3) descrevendo a energia cinética não aparece na equação (5) (o termo é igual a zero), e o termo de Coulomb referente ao núcleo-núcleo na equação (3) é uma constante. Este último, precisa ser adicionado à energia eletrônica, E^{el} , para se obter a energia total, E , do sistema.

$$E = E^{el} + \sum_{A < B}^{núcleo} \frac{Z_A Z_B}{R_{AB}} \quad (6)$$

A equação eletrônica de Schrödinger ainda é intratável e outras aproximações são necessárias. O mais óbvio é insistir que os elétrons se movem independentemente uns dos outros. Na prática, os elétrons individuais estão confinados a funções denominadas orbitais moleculares, cada uma das quais é determinada assumindo que o elétron está em movimento dentro de um campo médio (v^{HF}) resultante da presença dos demais elétrons, ou seja, cada movimento do elétron pode ser descrito por uma simples função (orbital molecular) que não depende explicitamente do movimento instantâneo dos outros elétrons.

Essa é a proposta do método de Hartree-Fock, uma metodologia aproximada para a determinação de funções de onda e energia no estado fundamental de sistemas quânticos de muitos corpos. A função de onda eletrônica antissimétrica é dada por um determinante de Slater, que é a função de onda antissimétrica mais simples que pode ser usada para descrever o estado fundamental de sistemas de N elétrons. A função de onda total é escrita como

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \cdot & \cdot & \cdot & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \cdot & \cdot & \cdot & \chi_N(x_2) \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \chi_1(x_N) & \chi_2(x_N) & \cdot & \cdot & \cdot & \chi_N(x_N) \end{vmatrix} \quad (7)$$

Aqui, χ_i são funções das coordenadas espaciais e spin de um único elétron, chamadas spin-orbitais moleculares. Admitindo que os χ_i 's são ortonormais, o fator $1/\sqrt{N!}$ é uma constante de normalização para Ψ .

Três importantes propriedades para um sistema composto de N elétrons são reunidas nesta função, sendo todas provenientes das propriedades dos determinantes. Trocando-se os elétrons da primeira e da segunda coluna, obtém-se:

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_2(x_1) & \chi_1(x_1) & \cdot & \cdot & \cdot & \chi_N(x_1) \\ \chi_2(x_2) & \chi_1(x_2) & \cdot & \cdot & \cdot & \chi_N(x_2) \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \chi_2(x_N) & \chi_1(x_N) & \cdot & \cdot & \cdot & \chi_N(x_N) \end{vmatrix} \quad (8)$$

Este determinante é igual ao original, porém com sinal oposto, ou seja, verifica-se que a função é antissimétrica para a troca de duas partículas:

$$\psi(1,2) = -\psi(2,1) \quad (9)$$

Escrevendo igualmente a primeira e a segunda coluna por exemplo, o determinante será igual a zero, portanto, não mais do que um elétron pode ocupar um orbital spin, impondo assim algebricamente o Princípio da Exclusão de Pauli.

Os elétrons são indistinguíveis e desta forma, o determinante garante que todos possam ocupar todos os orbitais com a mesma probabilidade.

O conjunto de orbitais que levam à menor energia são obtidos por um processo referido como campo autoconsistente (método SCF, do inglês, *Self-Consistent-Field*), que é capaz de obter soluções para as funções de onda do sistema a partir de uma função tentativa. Essa função tentativa é obtida a partir de uma estimativa inicial da função de onda, a qual é

aperfeiçoada até convergir. A ideia básica do método SCF é simples, através de um “chute” inicial dos orbitais de spin, pode-se calcular o campo médio experimentado por cada elétron, e em seguida, resolve-se a equação de Hartree-Fock para um novo conjunto de orbitais de spin. Usando estes novos orbitais de spin, tem-se um novo campo. Repete-se este procedimento até a autoconsistência ser alcançada. O arquétipo do método SCF é o método de Hartree-Fock. A equação de Hartree-Fock (HF) é uma equação de autovalor da forma:

$$f(i)\chi(x_i) = \epsilon\chi(x_i) \quad (10)$$

em que $f(i)$ pode ser escrito,

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{Z_A}{r_{iA}} + v^{HF}(i) \quad (11)$$

onde $f(i)$ é o operador de Fock de um elétron efetivo, x_i são as coordenadas spin e espacial do elétron i , χ_i é o orbital spin e $v^{HF}(i)$ é o potencial médio experimentado pelo elétron i devido à presença dos outros elétrons. A natureza do potencial médio v^{HF} depende da metodologia SCF.

O método de Hartree-Fock, em resumo, representa cada elétron por uma função de onda que possui dependência apenas nas coordenadas daquele elétron. A probabilidade de se encontrar um elétron em um dado ponto nas vizinhanças de um núcleo é determinada apenas pela posição em relação ao núcleo, mas não em relação aos outros elétrons. Há interação de um determinado elétron apenas com o campo médio dos outros elétrons, mas não há dependência mútua nas equações SCF, e não se considera as interações instantâneas entre esses elétrons, ou seja, os movimentos não estão correlacionados. O método de Hartree-Fock é o ponto de partida para todos os outros métodos quânticos *ab initio*⁶³⁻⁶⁵.

Os métodos *ab initio* tentam derivar informações pela resolução da equação de Schrödinger sem ajustar parâmetros pelos dados experimentais. Existem muitos métodos diferentes de aproximação para resolver a equação de Schrödinger, e normalmente o escolhido para determinado problema se dá pela comparação entre o seu desempenho e dados experimentais conhecidos. Desta forma, os dados experimentais orientam a seleção do modelo computacional mais propriamente do que participam diretamente do procedimento computacional.

Uma das aproximações inerentes a essencialmente todos os métodos *ab initio* é a introdução de um conjunto de bases, que é um conjunto de funções matemáticas a partir do

qual a função de onda é construída^{62,66}, ou seja, é o conjunto de funções usados para descrever a forma dos orbitais em um átomo. Orbitais moleculares e funções de onda são criados por combinações lineares das funções de base e funções angulares. O tipo de cálculo executado e o conjunto de bases escolhido são os dois fatores mais determinantes na precisão dos resultados⁶⁷.

A equação de Schrödinger é uma equação diferencial e sua resolução envolve a avaliação de um grande número de integrais. No caso de cálculos com métodos *ab initio*, o número de integrais cresce aproximadamente com a quarta potência do número de funções de base, chegando a alguns milhões mesmo para moléculas pequenas. Nos métodos quânticos semi-empíricos, a negligência de um grande número dessas integrais foi a solução adotada para economizar tempo de máquina e também reduzir a quantidade de memória necessária nos cálculos⁶¹. Tais métodos são denominados semi-empíricos porque parte de sua estrutura (o esqueleto formal) tem algum grau de fundamentação na teoria *ab initio*, e outra parte é empiricamente determinada. Na parametrização pode-se dispensar os dados experimentais de referência, empregando em vez disso, o resultado de cálculos *ab initio* de alto nível, e ainda assim chamar a metodologia produzida de semi-empírica, pois a otimização de parâmetros numéricos e a escolha de funções-tentativa não deixam de ser empirismos. As deficiências encontradas nos métodos semi-empíricos têm sua origem nas aproximações incorporadas em seu formalismo, assim como nas limitações inerentes ao problema da otimização de parâmetros. Também tem ocorrido a incorporação de erros em métodos semi-empíricos por meio do uso de informações experimentais incorretas⁶⁸.

Os métodos semi-empíricos são parametrizados para reproduzir vários resultados. Na maioria das vezes, geometria e energia (normalmente o calor de formação) são utilizados. Se inclui também, momentos de dipolo, calores de reação e potencial de ionização no conjunto de parametrização. Alguns métodos têm sido parametrizados para reproduzir uma propriedade específica, tais como, espectro eletrônico ou desvios químicos em ressonância magnética nuclear. Cálculos semi-empíricos podem calcular outras propriedades diferentes daquelas no conjunto de parametrização⁶⁷. Dentre os métodos semi-empíricos mais utilizados estão o MNDO (do inglês, *Modified Neglect of Diatomic Overlap*), AM1 (do inglês, *Austin Model 1*), e PM3 (do inglês, *Parametric Method 3*), dentre outros. A maioria dos métodos semi-empíricos utilizam um conjunto de bases pré-definido.

Neste trabalho, utiliza-se principalmente métodos da mecânica clássica e assim, a mecânica molecular será discutida mais detalhadamente a seguir.

3.3 Mecânica Molecular (MM)

O termo Mecânica Molecular foi criado na década de 70 para descrever a aplicação da mecânica clássica na determinação de estruturas moleculares em equilíbrio. A MM está baseada na aproximação de Born-Oppenheimer, caracterizada por considerar os movimentos nucleares e por tratar os elétrons indiretamente, diferindo dos métodos quânticos de orbitais moleculares, caracterizados por estudar a estrutura eletrônica, para uma posição nuclear fixa. Na MM é suposto que a densidade eletrônica pode ajustar-se instantaneamente a qualquer alteração na configuração geométrica dos núcleos, assumindo que os núcleos possuem movimento livre e que todas as interações nucleares são aditivas.

A superfície de Born-Oppenheimer, conhecida na Mecânica Molecular como superfície de energia potencial, é multidimensional e descreve a energia de uma molécula em termos das posições nucleares. Esta superfície possui mínimos locais em relação a uma ou algumas conformações moleculares estáveis e, em uma primeira aproximação, a molécula é descrita pela estrutura correspondente ao mínimo global da superfície potencial.

Na MM as moléculas são tratadas como uma coleção de átomos ligados que pode ser descrita por forças newtonianas, ou seja, são tratadas como uma coleção de partículas mantidas unidas por forças harmônicas ou elásticas. Estas forças são descritas pelas funções de energia potencial das contribuições estruturais, como os comprimentos de ligação (U_r), ângulos de ligação (U_θ), ângulos diedros (U_ϕ), interações não-ligadas como o potencial de Lennard-Jones ou van der Waals (U_d) e o potencial de Coulomb (U_c). O campo de força representa a energia potencial molecular em relação a uma geometria de referência, em um determinado sistema de coordenadas (equação 12):

$$U = \sum U(r) + \sum U(\theta) + \sum U(\phi) + \sum U(d) + \sum U(c) \quad (12)$$

O pressuposto fundamental do método da mecânica molecular é a transferibilidade dos parâmetros. Em outras palavras, a energia associada a um movimento molecular particular, por exemplo, o estiramento de uma ligação simples carbono-carbono, será a mesma de uma molécula para outra. Isso gera um cálculo muito simples que pode ser aplicado em sistemas moleculares muito grandes^{67,69,70}. A descrição dos termos da equação 12 serão detalhados a seguir.

3.3.1 Potencial de ligação

O potencial harmônico para o estiramento da ligação entre dois átomos, A e B, utilizando a lei de Hooke, contribui para a energia potencial total da molécula através da seguinte equação:

$$U(r_{AB}) = \frac{1}{2} k_{AB} (R_{AB} - R_{AB,eq})^2 \quad (13)$$

onde, o parâmetro k_{AB} representa a constante de força, R_{AB} o comprimento de ligação instantâneo e $R_{AB,eq}$ o comprimento de ligação em equilíbrio, sendo k_{AB} e $R_{AB,eq}$ parâmetros do campo de força⁶⁹.

3.3.2 Potencial angular

$U(\theta_{ABC})$ é o potencial harmônico que descreve a deformação angular entre três átomos A, B e C, onde existe uma ligação entre A e B e entre B e C. A aproximação harmônica se dá normalmente através da série de Taylor expandida até a segunda ordem sobre o ângulo de ligação,

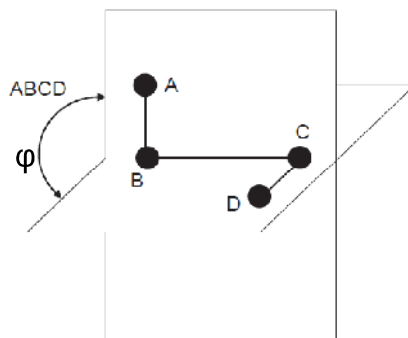
$$U(\theta_{ABC}) = \frac{1}{2} k_{ABC} (\theta_{ABC} - \theta_{ABC,eq})^2 \quad (14)$$

onde k_{ABC} é a constante de força, θ_{ABC} o ângulo de ligação instantâneo e $\theta_{ABC,eq}$ o ângulo de ligação em equilíbrio⁶⁶.

3.3.3 Potencial torsional

O potencial torsional ou diedral descreve parte da mudança de energia associada com a torção em torno da ligação B-C, em uma sequência de quatro átomos, A-B-C-D, onde A-B, B-C e C-D estão ligados. Tomando a ligação B-C, o ângulo torsional (φ) é definido como o ângulo formado pela ligação A-B e C-D como mostrado na Figura 15.

Figura 15 - Ângulo diedro.



Fonte:⁶⁹

Por interpretação, o ângulo torsional é periódico. Uma convenção óbvia é usar somente ângulo positivo, no caso o período de torção varia de 0 a 360°. No entanto, a energia mínima de muitas torções tem um arranjo antiperiplanar, isto é, $\varphi = \pi$. Assim, a convenção $-180^\circ \leq \varphi \leq 180^\circ$ é consideravelmente utilizada. Os diedros podem ser classificados em diedros próprios, onde se pode esperar uma rotação completa sobre a ligação B-C, e os diedros impróprios, onde essa rotação é limitada e portanto, são responsáveis por manter os grupos planares ou para evitar que as moléculas se sobreponham sobre suas imagens especulares. Por exemplo, se a ligação C-D fosse um fragmento C-H de um grupo metil, poderia se esperar uma rotação completa sobre a ligação B-C e uma simetria tripla na expressão de energia potencial. Por outro lado, uma união -CH-CH- no anel benzênico deve apresentar apenas uma flexão angular moderada em seu valor planar (ângulo zero).

O potencial diedral ou torsional é normalmente definido por:

$$U(\varphi) = \frac{U_0(\varphi)}{2} (1 - \cos(n(\varphi - \varphi_{eq}))) \quad (15)$$

onde $U_0(\varphi)$ é o potencial torsional inicial, n é o parâmetro de periodicidade, e φ_{eq} é o ângulo torsional no equilíbrio.

Os diedros impróprios também podem ser tratados da mesma forma que as deformações angulares, onde a contribuição para a energia potencial molecular é dada por:

$$U(\varphi_{ABCD}) = \frac{1}{2} k_{ABCD} (\varphi_{ABCD} - \varphi_{ABCD,eq})^2 \quad (16)$$

em que φ é o ângulo diedro^{62,66,69}.

3.3.4 Potencial de Lennard-Jones

As interações de van der Waals entre pares de átomos não-ligados covalentemente (i , j) são descritas pelo potencial de Lennard-Jones,

$$U_{LJ} = 4 \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (17)$$

onde, ϵ_{ij} é a profundidade do potencial entre a barreira repulsiva e atrativa, e σ_{ij} é a distância finita na qual o potencial inter-partícula é zero. Ambos são parâmetros ajustados experimentalmente ou por cálculos teóricos.

O potencial de Lennard-Jones é muito usado em campos de força cujo o alvo são sistemas moleculares muito grandes, como em biomoléculas. Em campos de força direcionados a sistemas de pequeno a médio porte, formas funcionais ligeiramente mais complicadas tendem a ser usadas^{62,71}.

3.3.5 Potencial de Coulomb

A outra parte da interação não-ligada é devido à redistribuição interna dos elétrons, criando partes positivas e negativas na molécula. Numa aproximação, isto pode ser modelado por cargas parciais para cada átomo. Alternativamente, pode-se atribuir um momento dipolar de ligação, sendo que estas duas descrições dão resultados similares (não idênticos). A interação entre cargas pontuais é dada pelo potencial de Coulomb,

$$U(c) = \frac{q_i q_j}{4 \pi \epsilon_0 \epsilon r_{ij}} \quad (18)$$

onde, r_{ij} é a distância entre os átomos i e j , q_i e q_j são suas cargas parciais, ϵ_0 a permissividade no vácuo e ϵ a constante dielétrica do meio. As cargas atômicas podem ser obtidas através de dados empíricos (cristalografia de raios-X, ressonância magnética nuclear, etc), mas são mais comumente ajustados ao potencial eletrostático através de cálculos mecânico-quânticos⁶⁶.

3.4 Ensemble Estatístico e Tratamento dos Sistemas Simulados

Uma grande variedade de propriedades termodinâmicas pode ser calculada das

simulações computacionais, em que uma comparação de valores experimentais e calculados para tais propriedades é uma importante forma na qual a precisão da simulação e o modelo de energia subjacente possa ser quantificado. Simulações podem também fornecer informações estruturais sobre as mudanças conformacionais nas moléculas e as distribuições das mesmas no sistema.

A maioria dos experimentos são realizados em amostras macroscópicas, contendo aproximadamente 10^{23} partículas. Cálculos, por outro lado, são realizados com relativamente poucas partículas, tipicamente $1-10^3$ até a 10^6 em casos especiais. O resultado macroscópico de uma medida experimental pode ser relacionado com as propriedades do sistema microscópico. A temperatura por exemplo, está relacionada com a energia cinética média das partículas:

$$\langle E_{cin} \rangle = \frac{3}{2} RT \quad (19)$$

A conexão entre as propriedades de um sistema microscópico e uma amostra macroscópica é fornecida pela mecânica estatística.

O estado de um sistema clássico pode ser completamente descrito especificando as posições e os momentos de todas as partículas. Sendo o espaço tridimensional, cada partícula tem associada a si, seis coordenadas – um sistema com N partículas é então caracterizado por $6N$ coordenadas (três coordenadas por átomo e três componentes do momento). O espaço dimensional $6N$ definido por essas coordenadas e momentos é chamado espaço de fase (τ) do sistema. Em qualquer instante no tempo, o sistema ocupa um ponto no espaço de fase.

$$X' = (x_1, y_1, z_1, p_{x,1}, p_{y,1}, p_{z,1}, x_2, y_2, z_2, p_{x,2}, p_{y,2}, p_{z,2}, \dots) \quad (20)$$

Para cada notação, as coordenadas de posição e de momento são definidas como,

$$q = (x_1, y_1, z_1, x_2, y_2, z_2, \dots) \quad (21)$$

$$p = (p_{x,1}, p_{y,1}, p_{z,1}, p_{x,2}, p_{y,2}, p_{z,2}, \dots) \quad (22)$$

permitindo escrever um ponto do espaço de fase (reordenado) como:

$$X = (q, p) \quad (23)$$

Portanto, um microestado fica definido pelos valores instantâneos das $3N$ coordenadas generalizadas q_i e os $3N$ momentos generalizados p_i e a maneira pela qual o sistema se move através do espaço de fase é governado pelas equações Hamiltonianas,

$$\begin{aligned}\frac{dq_i}{dt} &= \frac{\partial H(p, q)}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H(p, q)}{\partial q_i}\end{aligned}\tag{24}$$

onde $i = 1, \dots, 3N$, $H(p, q)$ é o Hamiltoniano do sistema e (p, q) representa um ponto ou vetor no espaço de fase com $6N$ componentes. Considerando um sistema clássico de N partículas, isolado, H não depende explicitamente do tempo, o sistema é conservativo e H é uma constante do movimento que corresponde a energia mecânica:

$$H(q, p) = E\tag{25}$$

A identidade anterior define uma superfície de energia no espaço de fase. A evolução do sistema conservativo é descrita por uma trajetória ou curva no espaço de fase sobre a superfície de energia. Como na Mecânica Clássica, cada condição inicial (p_0, q_0) determina de forma unívoca a evolução do sistema, trajetórias no espaço de fase nunca se cruzam.

Para uma dada energia E do sistema, existe um conjunto infinito de microestados. Definimos a função $\rho(p, q, t)$ como sendo a densidade de probabilidade de encontrar o sistema em um elemento de volume $dpdq$ no espaço de fase ao tempo t . O conjunto de pontos (p, q) cuja probabilidade ao tempo t é $\rho(p, q, t)dpdq$ formam um *ensemble* estatístico. Cada ponto representa uma cópia exata do sistema em um microestado diferente.

Alguns parâmetros macroscópicos podem ser mantidos constantes em conjunto, durante a simulação, como NpT , PVT , NVE ou μVT , onde N é o número de partículas no sistema, p a pressão, V o volume, T a temperatura, E é a energia total e μ é o potencial químico da substância. Os *ensembles* são caracterizados por estes conjuntos de parâmetros e consequentemente definem uma equação de estado para o sistema, permitindo que diferentes funções termodinâmicas possam ser mais convenientemente calculadas em um ou outro *ensemble*.

Como o espaço de fase abrange qualquer estado possível de um sistema, o valor médio de uma propriedade A no equilíbrio (isto é, o seu valor esperado) para um sistema com

temperatura, volume e o número de partículas constantes (NVT), por exemplo, pode ser escrito como uma integral sobre o espaço de fase,

$$\langle A \rangle = \iint A(q, p) P(q, p) dq dp \quad (26)$$

onde, P é a probabilidade de um ponto de fase particular. A partir da mecânica estatística, sabe-se que esta probabilidade depende da energia associada ao ponto de fase de acordo com

$$P(q, p) = Q^{-1} e^{-E(q, p)/k_B T} \quad (27)$$

onde E é a energia total (a soma da energia cinética e potencial em função de p e q , respectivamente), k_B é a constante de Boltzmann, T é a temperatura e Q é a função de partição do sistema

$$Q = \iint e^{-E(q, p)/k_B T} dq dp \quad (28)$$

a qual pode ser considerada como a constante de normalização de P .

Assumindo um sistema inicial em um ponto de fase “razoável” (ou seja, de baixa energia), sua evolução na conservação de energia ao longo do tempo (sua trajetória) parece provável para regiões relevantes da amostra no espaço de fase. Nesse caso, uma maneira razoável para computar a média de uma determinada propriedade, envolve, em simplesmente calcular o valor da propriedade periodicamente em tempos t_i e assumindo,

$$\langle A \rangle = \frac{1}{M} \sum_i^M A(t_i) \quad (29)$$

onde, M é o número de vezes que a propriedade é amostrada. No limite da amostragem, continuamente e seguindo o trajeto indefinidamente, essa equação se torna:

$$\langle A \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\tau) d\tau \quad (30)$$

A hipótese ergódica* assume a equação (30) como válida e independente da escolha de

* A dinâmica molecular se baseia na hipótese ergódica, que pressupõe que média obtida para um pequeno número de partículas durante um longo período de tempo é equivalente à média sobre um grande número de partículas sobre um curto intervalo de tempo.

t_0 . Para um gás no modelo de esfera rígida provou-se que a equação (26) e (30) são de fato, equivalentes. Tal prova não está disponível para sistemas mais realísticos, mas um grande número de evidências empíricas sugere que a hipótese ergódica é válida na maioria das simulações moleculares.

Para uma molécula única na aproximação do oscilador harmônico e rotor rígido, os estados de energia (quântico) são suficientemente regulares para permitir uma construção explícita da função de partição. Para uma coleção de muitas partículas interagindo (fase condensada), os estados energéticos relevantes são os que descrevem as vibrações, translações e rotações das moléculas em relação uma às outras. Para tais sistemas, os níveis de energia não são somente numerosos, mas também espaçados de modo irregular, o que impossibilita derivá-los diretamente a partir das quantidades moleculares. Por conseguinte, não é possível construir a função de partição explicitamente. É, no entanto, possível estimar derivadas de Q e diferenças em Q através de uma amostra representativa do sistema. As fases condensadas podem ser modeladas por condições periódicas de contorno e configurações geradas pelas técnicas de Dinâmica Molecular e Monte Carlo^{62,66,72-74}. Uma vez que o trabalho utilizou o procedimento de Dinâmica Molecular, o mesmo será descrito mais detalhadamente a seguir.

3.5 Dinâmica Molecular

As simulações por Dinâmica Molecular (MD, do inglês, *Molecular Dynamics*), estudam sistemas atômico-moleculares, em que os efeitos da temperatura não podem ser desprezados. Estão inclusos neste grupo, os sistemas químicos, físicos e biológicos constituídos por um grande conjunto de átomos e moléculas ($\sim 10^{23}$ partículas), onde as propriedades são governadas por processos que envolvem energias da ordem de $k_B T$, em que k_B é a constante de Boltzmann e T é a temperatura em Kelvin. A escala de energia envolvida nos processos e nas propriedades de interesse, numa faixa ao redor da temperatura ambiente, aproximadamente entre 100 a 800K, é aquela típica das interações entre átomos, íons e moléculas e das conformações de moléculas poliatômicas.

Por serem intrinsecamente termodinâmicos, a compreensão desses sistemas a nível atômico-molecular requer uma abordagem baseada na mecânica estatística, como citado anteriormente⁷³.

Uma simulação de dinâmica molecular requer um conjunto de coordenadas e velocidades iniciais, e um potencial de interação (função de energia). Para um intervalo de tempo curto, a interação pode ser considerada constante, permitindo que um conjunto atualizado de posições e velocidades possa ser estimado, no qual o ponto da nova interação

pode ser calculado. Tomando um número (grande) de intervalos de tempo (pequenos), o comportamento temporal do sistema pode ser obtido. Um intervalo de tempo típico de aproximadamente 10^{-15} segundos e uma simulação envolvendo 10^6 etapas, cobre “apenas” $\sim 10^{-9}$ segundos. Isto é substancialmente menor do que muitos fenômenos importantes, e uma vez que, o espaço de fase é enorme e o intervalo de tempo fundamental é pequeno, a simulação explorará somente a região próxima à configuração inicial, e desta forma, várias simulações diferentes, com condições iniciais diferentes são necessárias para estimar a estabilidade dos resultados.

Os núcleos são pesados o suficiente para que, em uma boa aproximação, se comportem como partículas clássicas e a dinâmica possa, desta forma, ser simulada resolvendo-se numericamente a equação de movimento de Newton,

$$F_i(t) = m_i a_i \quad (31)$$

que na forma diferencial pode ser escrita como,

$$\frac{d^2 x_i}{dt^2} = \frac{F_i(t)}{m_i} \quad (32)$$

em que,

$$F_i(t) = -\frac{\partial V(x_i)}{\partial x_i} \quad (33)$$

onde $F_i(t)$ é a força que atua sobre cada partícula do sistema em um instante de tempo t , e a_i é a aceleração do átomo i de massa m_i . Uma vez definido o campo de força, é possível calcular as forças que atuam sobre cada átomo, calculando-se a derivada primeira da energia potencial $V(x_i)$, obtida do campo de força escolhido, em relação às posições x_i desses átomos (equação 33). Esta equação gera diretamente a aceleração da partícula. A partir desta, integrando-se as equações de movimento, pode-se obter as velocidades, cuja integral, por sua vez, proporciona a mudança de posição do átomo. Com as novas posições e velocidades de cada partícula, obtêm-se as energias potencial e cinética do sistema. Aplicando-se sucessivamente esse procedimento, obtêm-se a trajetória, ou seja, o conjunto de posições e velocidades de cada partícula ao longo do tempo, a partir da qual serão obtidas as propriedades macroscópicas do

sistema^{66,71}.

No caso geral, a equação de movimento de cada partícula contém termos que dependem da posição e/ou velocidade de todas as outras partículas, de modo que a resolução analítica não é uma opção, e as equações de movimento são, ao invés disso, integradas numericamente no tempo. Esta é a primeira dentre muitas aproximações que devem ser feitas para que o método se torne capaz de tratar sistemas cujo tamanho e complexidade os aproxime daqueles estudados por químicos e biólogos⁷⁵.

3.6 Métodos de integração da equação de movimento

O método aplicado para integrar as equações de movimento é, naturalmente, uma aproximação. O parâmetro crucial é o tamanho do passo temporal, que influencia a ordem de grandeza do tempo de evolução do sistema que pode ser simulado no mesmo tempo de CPU (do inglês, *Central Processing Unit*), a ordem de grandeza do erro acumulado em cada passo, e os tipos de movimentos moleculares que podem ser investigados. Mais sutilmente, o método de integração pode influenciar a evolução do sistema resultando em um viés na exploração do espaço conformacional.

A integração das equações de movimento é realizada através de algoritmos baseados nos métodos das diferenças finitas, nos quais a integração é dividida em pequenos intervalos de tempo (passos de integração), Δt , permitindo simular os movimentos de maior frequência do sistema.

Dado um conjunto de partículas com posições r_i , as posições de um pequeno intervalo de tempo Δt posterior, são dados pela expansão de Taylor,

$$\begin{aligned} r_{i+1} &= r_i + \frac{\partial r}{\partial t}(\Delta t) + \frac{1}{2} \frac{\partial^2 r}{\partial t^2}(\Delta t)^2 + \frac{1}{6} \frac{\partial^3 r}{\partial t^3}(\Delta t)^3 + \dots \\ r_{i+1} &= r_i + v_i(\Delta t) + \frac{1}{2} a_i(\Delta t)^2 + \frac{1}{6} b_i(\Delta t)^3 + \dots \end{aligned} \quad (34)$$

As velocidades v_i , são as primeiras derivadas das posições em relação ao tempo (dr/dt) no tempo t_i , as acelerações a_i , são as segundas derivadas (d^2r/dt^2) no tempo t_i , as hiperacelerações b_i , são as terceiras derivadas, etc. As posições de pequenos intervalos de tempo Δt anteriores são derivadas da equação (34) substituindo Δt por $-\Delta t$.

$$r_{i-1} = r_i - v_i(\Delta t) + \frac{1}{2} a_i(\Delta t)^2 - \frac{1}{6} b_i(\Delta t)^3 + \dots \quad (35)$$

A adição das equações (34) e (35) gera um modelo para prever a posição (r_{i+1}) em um intervalo de tempo Δt posterior a partir das posições atuais e anteriores (r_i e r_{i-1}) e da aceleração imediata a_i . Esta última pode ser calculada a partir da força, e portanto, equivalentemente do potencial.

$$\begin{aligned} r_{i+1} &= (2r_i - r_{i-1}) + a_i (\Delta t)^2 + \dots \\ a_i &= \frac{F_i}{m_i} = -\frac{1}{m_i} \frac{dV}{dr_i} \end{aligned} \quad (36)$$

Este é o algoritmo de Verlet⁷⁶ para resolver numericamente a equação de Newton, em que o termo envolvendo a mudança na aceleração (b_i) desaparece, ou seja, a equação é correta para a terceira ordem em Δt . No ponto inicial, as posições anteriores não estão disponíveis, mas podem ser estimadas a partir de uma aproximação de primeira ordem da equação (34).

$$r_{i-1} = r_0 - v_0 \Delta t \quad (37)$$

A cada intervalo de tempo, a aceleração deve ser avaliada, equação (36), na qual então permite que as posições atômicas sejam propagadas no tempo e assim gerando a trajetória.

O algoritmo de Verlet tem a desvantagem numérica que as novas posições são obtidas adicionando um termo proporcional à Δt^2 para uma diferença nas posições ($2r_i - r_{i-1}$). Uma vez que, (Δt) é um número pequeno e $(2r_i - r_{i-1})$ é uma diferença entre dois números grandes, isto pode conduzir a erros de truncagem devido à precisão finita. Além disso, o algoritmo de Verlet tem a desvantagem de que as velocidades não aparecem explicitamente, na qual é um problema em conexão com *ensembles* gerados com temperatura constante.

O aspecto numérico e a falta das velocidades explícitas no algoritmo de Verlet pode ser remediado pelo algoritmo *leap-frog*⁷⁷. Desenvolvendo expansões análogas às equações (34) e (35) com a metade de um intervalo de tempo seguido por subtração gera,

$$r_{i+1} = r_i + v_{i+\frac{1}{2}} \Delta t \quad (38)$$

A velocidade é obtida pela expansão análoga para dar a equação (39).

$$v_{i+\frac{1}{2}} = v_{i-\frac{1}{2}} + a_i \Delta t \quad (39)$$

As equações (38) e (39) definem o algoritmo de *leap-frog*, e se observa que as atualizações das posições e velocidades estão fora de fase pela metade do intervalo de tempo, ou seja, as velocidades são atualizadas em tempos intermediários àqueles em que são atualizadas as posições. Em termos de precisão teórica ele é também de terceira ordem, como o algoritmo de Verlet, mas a precisão numérica é melhor. Além disso, a velocidade aparece diretamente, o que facilita um acoplamento com um banho térmico externo. A desvantagem é que as posições e velocidades não são conhecidas ao mesmo tempo, por eles estarem sempre fora de fase pela metade do intervalo de tempo. A última anormalidade pode ser removida pelo algoritmo de velocidade de Verlet, onde as equações utilizadas para propagar os átomos são dadas na equação (40)^{66,71,75}.

$$\begin{aligned} r_{i+1} &= r_i + v_i \Delta t + \frac{1}{2} a_i \Delta t^2 \\ v_{i+1} &= v_i + \frac{1}{2} (a_i + a_{i+1}) \Delta t \end{aligned} \quad (40)$$

O intervalo de tempo empregado é um parâmetro de controle importante para a simulação. O intervalo de tempo máximo que pode ser utilizado é determinado pela velocidade do processo mais rápido no sistema, ou seja, tipicamente uma ordem de magnitude menor do que o processo mais rápido. Movimentos moleculares (rotações e vibrações) ocorrem caracteristicamente em frequências que variam entre 10^{11} - 10^{14} s⁻¹ (correspondente a comprimentos de ondas de 3-3300 cm⁻¹), e intervalos de tempo da ordem de femtossegundos (10^{-15} s) ou menos são necessários para modelar tais movimentos com precisão suficiente.

Infelizmente, muitos fenômenos de interesse ocorrem em uma escala de tempo substancialmente grande: enovelamento de proteínas e reações químicas, por exemplo, ocorrem em ordens de milissegundos ou segundos. Além disso, uma única trajetória pode não ser adequada para representar a dinâmica, exigindo assim, que muitas análises devam ser feitas com diferentes condições de partida (posições e velocidades) e que suas médias sejam adequadamente obtidas.

Para moléculas, os processos mais rápidos são os estiramentos das ligações, especialmente aqueles envolvendo os hidrogênios. Estes graus de liberdade, no entanto, têm relativamente pouca influência sobre várias propriedades. Com isso, se torna vantajoso “congelar” todos os comprimentos de ligação envolvendo os átomos de hidrogênio, permitindo se utilizar intervalos de tempo mais longos, e consequentemente tempos de simulações maiores, obtidos com um mesmo custo computacional. Como todos os átomos se

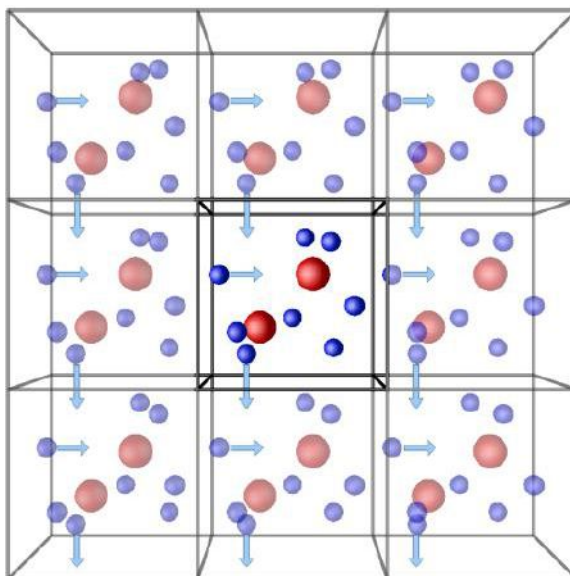
movem individualmente de acordo com a equação de Newton, as restrições devem ser aplicadas para manter os comprimentos de ligação fixos. Isto é feito normalmente pelos algoritmos *SHAKE*⁷⁸ (Verlet) ou *RATTLE*⁷⁹ (velocidade de Verlet), onde a distância constrangida é incorporada pelo método de multiplicadores indeterminados de Lagrange. Primeiramente, é permitido que os átomos se movam sobre a influência de forças, e subsequentemente são forçados a obedecer às restrições, tomando alguns passos sequenciais através de todas as variáveis⁶⁶.

A vantagem dos modelos com restrições é clara, porém esses modelos utilizados para reproduzir a dinâmica de sistemas sem restrições possui limitações. Van Gunsteren e Karplus⁸⁰ mostraram através de simulações no vácuo, da proteína bovina tripsina, inibidora pancreática, (BPTI do inglês, *bovine pancreatic trypsin inhibitor*) que o uso de comprimentos de ligação fixados não altera significativamente as propriedades dinâmicas do sistema, enquanto que fixando os ângulos de ligação, ocorrem alterações significativas⁶⁰.

3.7 Condições periódicas de contorno

Em sistemas reais macroscópicos, o número de partículas é da ordem do número de Avogadro, onde a quantidade de partículas na superfície é relativamente pequena. Em sistemas simulados através da Dinâmica Molecular, na ordem de 10^2 - 10^6 partículas, o número relativo de átomos e moléculas na superfície é muito grande, causando efeitos de superfície indesejáveis. Para minimizar esse efeito utiliza-se condições periódicas de contorno. Os átomos do sistema a ser simulado são colocados em uma caixa, normalmente com geometria cúbica, e então cópias idênticas desta caixa são replicadas em todas as direções do espaço, Figura 16. Durante as simulações, quando um átomo move-se na caixa original, sua imagem periódica, em uma das caixas imagem, move-se da mesma maneira. Se um átomo sai da caixa original, sua imagem entra pela face oposta da caixa imagem, com a mesma velocidade. Assim, o número total de átomos na caixa central e no sistema como um todo é conservado. Isto significa que o modelo resultante se torna quase periódico, com uma periodicidade igual às dimensões da caixa^{66,71}.

Figura 16 - Representação das Condições Periódicas de Contorno em três dimensões.

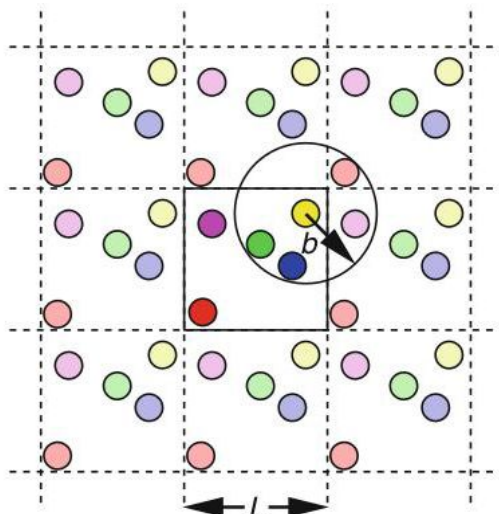


Fonte: <http://isaacs.sourceforge.net/phys/pbc.html> acessado em 10/10/2015.

3.8 Tratamento das interações de curto e longo alcance

Para moléculas grandes, existe muito mais interações não-ligadas do que interações ligadas. Desta forma, com a implementação das condições periódicas de contorno as interações não-ligadas se estendem além do limite da caixa englobando as interações com as partículas das caixas imagem, tornando infinito os termos da soma dos potenciais, e na prática, portanto, se torna necessário a realização de aproximações. Para evitar a interação da partícula com sua própria imagem ou com a mesma partícula duas vezes e para o tratamento das interações de curto alcance, onde o potencial de interação intermolecular decai rapidamente com o aumento da distância entre as partículas $V \sim r^{-6}$ (interações de van der Waals – potencial de Lennard-Jones), emprega-se o procedimento denominado convenção da imagem mínima. Nesse caso, é empregado o truncamento do potencial em um raio de corte esférico b , geralmente da ordem de $L/2$ (L é o comprimento da célula de simulação), Figura 17.

Figura 17 - Representação do raio de corte esférico em duas dimensões.

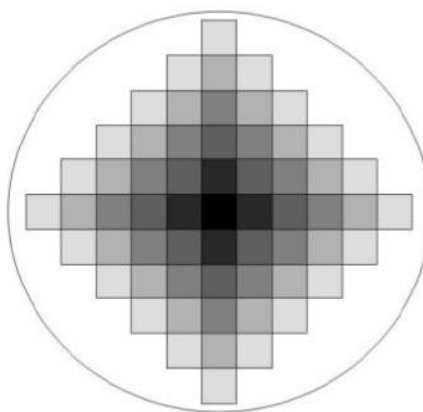


Fonte:⁶⁰.

Para o tratamento das interações de longo alcance (interações eletrostáticas – potencial de Coulomb), em que o potencial decai com a distância que é menor do que r^{-d} (d é a dimensionalidade do sistema), o truncamento abrupto do potencial produz efeitos que, em muitos casos, comprometem severamente os resultados das simulações. Vários métodos são utilizados para tratar as forças de longo alcance, dentre eles a soma de Ewald⁸¹ e aproximação do campo de reação⁸².

Na soma de Ewald, uma partícula interage com todas as outras partículas na caixa de simulação e com todas as suas imagens num arranjo infinito de células periódicas, sendo que no limite, a geometria do arranjo das células é considerada esférica, Figura 18.

Figura 18 - Construção de um sistema de células periódicas no método de Ewald.



Fonte:⁷².

A contribuição da interação carga-carga para a energia potencial atribuída a todos os pares de cargas na caixa de simulação central pode ser escrita como,

$$V = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{4 \pi \epsilon_0 r_{ij}} \quad (41)$$

onde r_{ij} é a distância mínima entre as cargas i e j , e ϵ_0 é a permissividade no vácuo. Já a contribuição da interação carga-carga entre as cargas na caixa central e todas as imagens de todas as partículas das seis caixas vizinhas é dado por:

$$V = \frac{1}{2} \sum_{n_{caixa}=1}^6 \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{4 \pi \epsilon_0 |r_{ij} + r_{caixa}|} \quad (42)$$

Existe portanto, uma contribuição para a energia total a partir das interações na caixa central somado às interações entre a caixa central e todas as suas caixas imagens^{69,71,72}.

O método de Ewald é largamente usado em sistemas altamente polares ou carregados. Existem variantes do método original de Ewald como o método PME (do inglês, *Particle Mesh Ewald*)⁸³ que utiliza procedimentos de interpolação por *splines* para tratar os termos no espaço recíproco. Por ser muito eficiente em termos de precisão e velocidade de processamento, o método PME encontra-se implementado em vários pacotes comerciais de simulação e de domínio público⁷³.

3.9 Modelos de solvatação contínua

Quando um soluto é imerso em um solvente, sua distribuição de carga interage com aquela do solvente, portanto, em vez de representar a distribuição de carga do solvente explicitamente, substitui-se por um campo elétrico contínuo que representa a média estatística sobre todos os graus de liberdade do solvente em equilíbrio térmico.

Todos os modelos de solvatação contínua são baseados na equação de Poisson da eletrostática clássica, para expressar o potencial eletrostático em função da densidade de carga e da constante dielétrica. A equação de Poisson é uma equação diferencial de segunda ordem escrita como,

$$\nabla \epsilon(r) \nabla \varphi(r) = -4 \pi \rho(r) \quad (43)$$

Neste caso, a constante dielétrica pode depender da posição. A equação de Poisson,

válida para situações em que o meio dielétrico circundante responde de forma linear à incorporação de carga, ou seja, a constante dielétrica é independente da posição (isto é, de fato uma constante), é escrita como,

$$\nabla^2 \varphi(r) = \frac{-4\pi\rho(r)}{\varepsilon} \quad (44)$$

onde ε é a constante dielétrica do meio. Se a densidade de carga é uma carga pontual, a solução da equação (44) se reduz para a interação de Coulomb. A equação (43) pode ser usada para descrever, por exemplo, a solvatação de uma proteína em água, onde a região da proteína é considerada como tendo uma constante dielétrica pequena ($2 < \varepsilon < 5$) enquanto que o solvente apresenta uma constante dielétrica alta ($\varepsilon = 78$).

A equação de Poisson pode ser modificada levando em consideração uma distribuição de Boltzmann (térmica) de íons no solvente. Os íons negativos se acumulam onde o potencial é positivo, e vice-versa, sujeitos à flutuação térmica. As densidades de carga de uma coleção de íons com cargas q e $-q$ e concentração c são dadas por:

$$\begin{aligned} \rho_+ &= qce^{-q\varphi/kT} \\ \rho_- &= qce^{-q\varphi/kT} \end{aligned} \quad (45)$$

A adição destas contribuições à equação (43) conduz à equação de Poisson-Boltzmann,

$$\begin{aligned} \nabla\varepsilon(r)\nabla\varphi(r) - \varepsilon(r)\lambda(r)\kappa^2 \frac{k_B T}{q} \sinh\left[\frac{q\varphi(r)}{k_B T}\right] &= -4\pi\rho(r) \\ \kappa^2 &= \frac{8\pi q^2 I}{\varepsilon k_B T} \end{aligned} \quad (46)$$

onde λ é uma função de troca simples na qual é zero em regiões inacessíveis ao eletrólito e num caso contrário, κ^2 é o parâmetro de Debye-Hückel sendo I a força iônica da solução eletrólito. O termo $\sinh[q\varphi(r)/K_B T]$ é aplicado somente para região correspondente ao solvente, ou seja, para r fora da cavidade. Uma vez que $q\varphi/kT$ é adimensional, a equação anterior é, muitas vezes escrita em termos de um potencial reduzido u ,

$$\nabla\varepsilon(r)\nabla\varphi(r) - \varepsilon(r)\lambda(r)\kappa^2 \sinh(u(r)) = -4\pi\rho(r) \quad (47)$$

Se o potencial é suficientemente pequeno (isto é, o soluto não é fortemente carregado), a função $\sinh(x)$ pode ser expandida em uma série de Taylor, $\sinh(x) \approx x + x^3/6 + \dots$.

Mantendo somente o primeiro termo, tem-se a equação de Poisson-Boltzmann linearizada,

$$\nabla \varepsilon(r) \nabla \varphi(r) - \varepsilon(r) \lambda(r) \kappa^2 u(r) = -4 \pi \rho(r) \quad (48)$$

Todas essas equações (43) à (48) são equações diferenciais que devem ser resolvidas numericamente, e os resultados dão informações sobre o potencial eletrostático em qualquer ponto do espaço. O potencial pode ser mapeado sobre a superfície do soluto onde pode sugerir regiões para interação com outras moléculas polares^{62,66}.

Um método de solvatação contínua bastante utilizado no caso em que as somas de Ewald e suas variantes não são uma opção viável, é o método do campo de reação, o qual simplesmente reescala as interações eletrostáticas de longo alcance por uma constante dielétrica, fornecida pelo usuário, que represente a constante dielétrica do solvente. Neste método, uma esfera é construída em volta da molécula com um raio igual à distância do raio de corte. A interação das moléculas que estão dentro da esfera é calculada explicitamente. Para isto, é adicionado uma energia de interação sobre o ambiente fora da esfera, na qual é modelada como um meio de constante dielétrica homogênea ε_{rf} , Figura 19, em que o campo de reação no centro da cavidade experimentado pela molécula a é dado por:

$$E_a = \frac{2(\varepsilon_{rf} - 1)}{\varepsilon_{rf} + 1} \left(\frac{1}{r_c^3} \right) \sum_{b; r_{ab} \leq r_c} \mu_b \quad (49)$$

onde $\sum \mu_b$ é o momento de dipolo total de todas as moléculas dentro da esfera de raio de corte r_c . Assumindo que (i), a carga total em cada cavidade, é igual a zero, e (ii) cada molécula a é suficientemente pequena, de tal modo que o conjunto de átomos na cavidade da molécula a é igual aos átomos da esfera de raio de corte do átomo i , para cada i da molécula a , a energia do campo de reação (E^{RF}) do sistema de M moléculas é:

$$\begin{aligned} E^{RF}(r_1, \dots, r_N) &\equiv -\frac{1}{2} \sum_{a=1}^M (\mu \nabla E_a) \\ E^{RF}(r_1, \dots, r_N) &= \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j(\neq i) \\ r_{ij} < r_c}} q_i q_j \frac{\varepsilon_{rf} - 1}{2\varepsilon_{rf} + 1} \frac{r_{ij}^2}{r_c^3} \end{aligned} \quad (50)$$

Assim a energia total se torna:

$$E_{RF}(r_1, \dots, r_N) = \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j(\neq i) \\ r_{ij} \leq r_c}} \frac{q_i q_j}{r_{ij}} + E^{RF}(r_1, \dots, r_N) \quad (51)$$

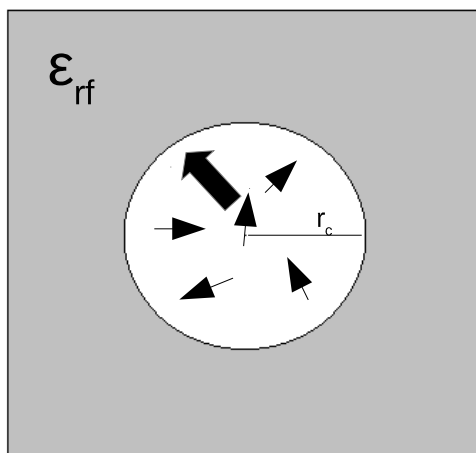
$$E_{RF}(r_1, \dots, r_N) = \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j(\neq i) \\ r_{ij} \leq r_c}} q_i q_j V_{RF}(r_{ij})$$

com,

$$V_{RF}(r) \equiv \frac{1}{r} \left[1 + \frac{\epsilon_{rf} - 1}{2\epsilon_{rf} + 1} \left(\frac{r}{r_c} \right)^3 \right] \quad (52)$$

O método do campo de reação tem a vantagem de ser conceitualmente simples, relativamente fácil de se implementar e eficiente computacionalmente^{62,72,84}.

Figura 19 - Método do campo de reação. A seta com a haste sombreada representa a soma dos dipolos das moléculas dentro da esfera de raio de corte r_c .



Fonte: Autor

3.10 Dinâmica molecular de sistemas aquosos

A água desempenha um papel crucial na estrutura e termodinâmica de macromoléculas. Em particular, a estabilidade e funcionalidade das proteínas e ácidos nucleicos são ditadas por efeitos específicos bem como não específicos do solvente. Enzimas, por exemplo, requerem um certo nível de hidratação antes que possam desenvolver sua função biológica. A maioria das funções celulares são governadas por mudanças no ambiente do solvente, incluindo variação de pH e atividades iônicas, bem como diferenças nas concentrações de soluto entre os compartimentos celulares e subcelulares. Portanto, é

importante que as simulações biomoleculares devam tentar recriar o ambiente aquoso de biomoléculas de forma mais exata possível, incluindo os efeitos da água, implícita ou explicitamente⁸⁵⁻⁸⁷.

Na solvatação implícita, em um dos modelos mais populares, os graus de liberdade do solvente não são considerados explicitamente e, desta forma, o solvente é tratado como um meio contínuo tendo as propriedades médias do solvente real. Este modelo de campo de reação foi discutido na seção anterior e para moléculas de forma arbitrária, as equações de Poisson-Boltzmann que descrevem as interações eletrostáticas são resolvidas por métodos de diferenciação finita. Porém, para sistemas grandes, como proteínas e ácidos nucleicos, métodos de simplificação como a aproximação generalizada de Born são utilizadas.

Neste modelo, a contribuição eletrostática para a energia livre de solvatação é:

$$\Delta G_{pol} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w} \right) \sum_{i,j} \frac{q_i q_j}{f_{GB}} \quad (53)$$

onde q_i e q_j são as cargas parciais, ϵ_w é a constante dielétrica do solvente, e f_{GB} é a função que interpola entre um “raio de Born efetivo” R_i , quando a distância r_{ij} entre os átomos é curta, e r_{ij} quando há grandes distâncias⁸⁸.

No modelo de solvatação explícita, o potencial pode ser parametrizado para reproduzir propriedades macroscópicas, para uma fase escolhida sob determinadas condições. Esta abordagem empírica deu origem à família dos potenciais TIP (do inglês, *Transferable Intermolecular Potentials*) desenvolvido por Jorgensen e colaboradores: TIP3P, TIP4P⁸⁹ e TIP5P⁹⁰. Tais modelos usam três, quatro e cinco sítios de interação (Figura 20), respectivamente, sendo todos estes modelos tratados como rígidos. São modelos mais simples que se baseiam nas interações não ligadas, uma vez que, as interações ligadas são tratadas implicitamente. A função potencial entre duas moléculas de água a e b para os modelos da família TIP n P é representada por:

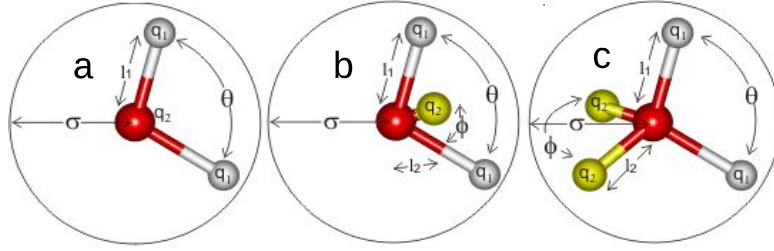
$$V_{ab} = \sum_{ij} \frac{q_i q_j e^2}{r_{ij}} + 4 \epsilon_0 \left[\left(\frac{\sigma_0}{r_{oo}} \right)^{12} - \left(\frac{\sigma_0}{r_{oo}} \right)^6 \right] \quad (54)$$

onde i e j são os sítios de carga em a e b , respectivamente, separadas por uma distância r_{ij} , ϵ_0 e σ_0 são os parâmetros de van der Waals entre os dois sítios de oxigênio, os quais são separados por r_{oo} .

Outro modelo de interação para a água comumente utilizado é SPC⁹¹ (do inglês,

Simple Point Charge), no qual as cargas nos três sítios de interação do hidrogênio carregados positivamente são balanceados por uma carga negativa apropriada no oxigênio, Figura 20.

Figura 20 - Representação dos modelos rígidos na solvatação explícita; (a) de três sítios de interação – SPC, TIP3P; (b) de quatro sítios – TIP4P; (c) de cinco sítios – TIP5P; l_1 e l_2 representam as distâncias entre o átomo hidrogênio e oxigênio; q_1 e q_2 representam as cargas atômicas do hidrogênio e oxigênio, respectivamente; θ e Φ são os ângulos de ligação e σ é a distância para o potencial zero.



Fonte: http://www1.lsbu.ac.uk/water/water_models.html, acessado em 29/09/2015.

3.11 Acoplamento de temperatura e pressão

Uma simulação à temperatura constante pode ser necessária quando se deseja determinar como a conduta de um sistema varia com a temperatura, tal como num processo de enovelamento de proteína. A temperatura do sistema está relacionada com a média da energia cinética, a qual para um sistema não-constrangido é dado por,

$$\langle E_{cin} \rangle_{NVT} = \frac{3}{2} k_B RT \quad (55)$$

uma maneira de alterar a temperatura do sistema é então, dimensionar as velocidades. Se a temperatura no tempo t é $T(t)$ e as velocidades são multiplicadas por um fator λ , então a variação de temperatura associada pode ser calculada como segue:

$$\Delta T = \frac{1}{2} \sum_{i=1}^N \frac{2}{3} \frac{m_i (\lambda v_i)^2}{N k_B} - \frac{1}{2} \sum_{i=1}^N \frac{2}{3} \frac{m_i v_i^2}{N k_B} \quad (56)$$

$$\Delta T = (\lambda^2 - 1) T(t) \quad (57)$$

$$\lambda = \sqrt{\frac{T_{desejada}}{T(t)}} \quad (58)$$

A maneira mais simples de calcular a temperatura é, então, multiplicar as velocidades em cada intervalo de tempo por um fator $\lambda = \sqrt{T_{(desejada)}/T_{(atual)}}$, onde $T_{(atual)}$ é a temperatura atual calculada da energia cinética e $T_{(desejada)}$ é a temperatura escolhida.

Uma maneira alternativa de controlar a temperatura é acoplar o sistema a um banho térmico externo fixado na temperatura desejada, processo conhecido como termostato de Berendsen⁹². O banho atua como uma fonte de energia térmica fornecendo ou removendo calor do sistema quando apropriado. As velocidades são reescaladas a cada passo de integração, a fim de ajustar a energia cinética do sistema até a temperatura selecionada, tal que a taxa de variação da temperatura é proporcional à diferença entre as temperaturas do banho e do sistema,

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_{banho} - T(t)) \quad (59)$$

onde τ é o parâmetro de acoplamento, cuja magnitude determina o quão firmemente o banho e o sistema estão acoplados. Esse método gera um decaimento exponencial do sistema rumo à temperatura desejada. A mudança na temperatura entre os sucessivos intervalos de tempo é:

$$\Delta T = \frac{\delta t}{\tau} (T_{banho} - T(t)) \quad (60)$$

O fator de reescalamento para as velocidades é, portanto,

$$\lambda^2 = 1 + \frac{\delta t}{\tau} \left(\frac{T_{banho}}{T(t)} - 1 \right) \quad (61)$$

se τ é grande, então o acoplamento será fraco. Se τ é pequeno, o acoplamento será forte e quando o parâmetro de acoplamento é igual ao intervalo de tempo ($\tau = \delta t$), então o algoritmo é equivalente ao método simples de reescalamento de velocidade. Outros métodos mais sofisticados como os termostatos de Nosé-Hoover^{93,94} e Nosé- Poincaré⁹⁵ também são utilizados.

Muitas medidas experimentais são realizadas sob condições de temperatura e pressão constante, e portanto, simulações em *ensembles* isotérmico-isobáricos são diretamente mais relevantes para dados experimentais. De maneira análoga à temperatura, pode-se controlar a pressão substituindo as variáveis T por P , e as velocidades dos átomos pelas coordenadas

atômicas nas equações mostradas anteriormente. Dentre os barostatos destacam-se o de Berendsen⁹² e o de Parrinello-Rahman⁹⁶.

3.12 Minimização de energia

A maneira na qual a energia varia em relação às coordenadas (sejam internas ou cartesianas) é usualmente referida como superfície de energia potencial. Na modelagem molecular se está, especialmente interessado nos pontos de mínimo da superfície de energia. Arranjos dos átomos com energia mínima correspondem a estados estáveis do sistema, onde qualquer movimento fora do mínimo gera uma configuração com energia mais alta. Pode haver um número muito grande de mínimos na superfície de energia em que o de menor energia é conhecido como energia mínima global, enquanto todos os outros são descritos como mínimos locais. Para identificar estas geometrias do sistema que correspondem aos pontos de mínimo na superfície de energia se utiliza algoritmos de minimização.

Seja uma função f na qual dependa de uma ou mais variáveis independentes x_1, x_2, \dots, x_i . Encontrar os valores destas variáveis onde f tenha um valor mínimo é o objetivo destes algoritmos, os quais podem ser classificados em dois grupos. Os que utilizam derivadas da energia com relação às coordenadas e os que não utilizam. As derivadas podem ser úteis, uma vez que, podem fornecer informações sobre a forma da superfície de energia, e se utilizadas adequadamente, podem melhorar significativamente a eficiência com a qual o mínimo é localizado, e por isso são utilizadas pelos métodos de minimização mais populares.

No ponto mínimo, a primeira derivada da função em relação a cada uma das variáveis é zero e as segundas derivadas são sempre positivas:

$$\frac{\partial f}{\partial x_i} = 0; \frac{\partial^2 f}{\partial x_i^2} > 0 \quad (62)$$

Dentre os métodos que utilizam derivadas estão o método de *Steepest Descent*⁹⁷, o método do *gradiente conjugado*⁹⁸ e o método de *Newton-Raphson*⁹⁹. O método utilizado neste trabalho foi o de *Steepest Descent* e este será descrito em seguida.

O método de *Steepest Descent* é um algoritmo de otimização de primeira ordem que, embora não seja certamente o mais eficiente para minimização, é robusto e de fácil implementação. Para encontrar um mínimo (local) de uma função usa-se um esquema iterativo, onde em cada passo se toma a direção (negativa) do gradiente ∇ , que corresponde à direção de declive máximo. Dada uma coordenada inicial x_o , o método computa uma

sequência de iterações x_k , onde

$$x_{k+1} = x_k - t_k \nabla f(x_k), \quad k=0,1,2,\dots, \quad (63)$$

onde $t_k > 0$ minimiza a função

$$\phi_k(t) = f(x_k - t \nabla f(x_k)) \quad (64)$$

Por natureza, o método de *Steepest Descent* só pode localizar funções mínimas. A vantagem é que o algoritmo é muito simples e requer apenas o armazenamento do vetor gradiente. Além disso, é um dos poucos métodos que garante a diminuição do valor da função. Seu principal uso é “relaxar” rapidamente uma configuração inicial pobre (de elevada energia)^{66,72,100}.

4 Metodologia

4.1 Sistemas simulados

A estrutura inicial da enzima EPSPs simulada neste trabalho, foi obtida do banco de dados de proteínas (RCSC-PDB (do inglês, *Research Collaboratory for Structural Bioinformatics – Protein Data Bank*)¹⁰¹ cujo o código do modelo cristalográfico no banco é 1G6S¹⁰. Simulação de Dinâmica Molecular (DM) através do pacote computacional GROMACS 4.5.3¹⁰², em conjunto com o campo de força OPLS-AA (do inglês, *Optimized Potentials for Liquid Simulations – all-atom*)¹⁰³ foi utilizada para obter a dinâmica estrutural da enzima EPSPs em solução aquosa, sem a presença do herbicida glifosato (GPJ) e o substrato Shikimato-3-Fosfato (S3P).

A enzima EPSPs com estrutura fechada, possuindo 6480 átomos foi colocada numa caixa cúbica de simulação e então solvatada explicitamente segundo o modelo SPC, e neutralizada até atingir um sistema com um total de 127995 átomos.

Em seguida foi realizado a minimização de energia do sistema num tempo de 10 ps, utilizando o algoritmo *Steepest Descent*, sendo o tratamento das interações eletrostáticas obtido através do método PME.

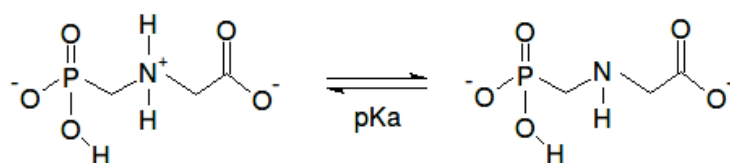
A temperatura do sistema foi aumentada realizando-se sucessivas simulações por dinâmica molecular com tempos de 100 ps à 50, 150 e 298 K, em que durante este processo, as posições dos átomos da proteína, com exceção dos hidrogênios, são restringidas aplicando-se uma constante de força de $1,0 \cdot 10^{-3} \text{ KJ.mol}^{-1}.\text{nm}^{-2}$, enquanto as moléculas de água ficaram livres para acomodar-se de maneira otimizada ao longo de sua superfície.

A dinâmica molecular do sistema foi realizada num intervalo de tempo de integração de 2 fs totalizando um tempo de 50 ns em um *ensemble* NPT (número de partículas, pressão e temperatura constantes), utilizando o algoritmo *leap-frog* para integrar as equações de movimento de Newton a uma pressão de 1 bar e temperatura de 298 K, utilizando termostato *v-rescale* e barostato de Berendsen. As restrições dos estiramentos lineares e angulares além do comprimento de ligação dos átomos de hidrogênio foram realizadas pelo algoritmo LINCS (do inglês, *Linear Constraint Solver*)¹⁰⁴.

4.2 Parametrização do herbicida glifosato

As duas formas predominantes da estrutura da molécula de glifosato em um $pK_a = 5,69$ são mostradas na Figura 21. A estrutura protonada, com carga (1-) foi utilizada para a parametrização neste trabalho.

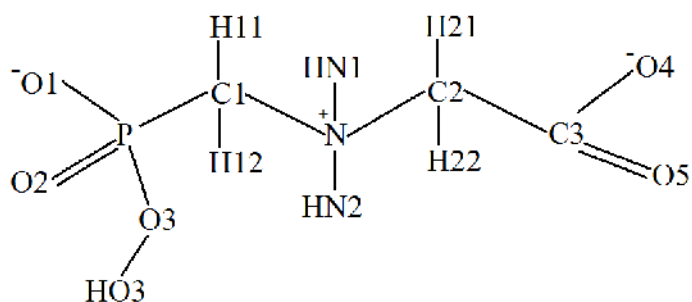
Figura 21 - Representação esquemática da dissociação do herbicida glifosato, em pKa = 5,69.



Fonte: Autor.

Partindo da forma funcional do campo de força OPLS-AA para estiramento de ligação, deformação angular e diedral (Tabela 1), foi utilizado um novo conjunto de cargas para os átomos do herbicida glifosato e um potencial de interação torsional desenvolvido para os diedros próprios entre os átomos HO3-O3-P-C1 e entre os átomos P-C1-N-HN1 (Figura 22), uma vez que, estes diedros não são encontrados no campo de força OPLS-AA.

Figura 22 - Representação da molécula glifosato: Diedros parametrizados HO3-O3-P-C1 e P-C1-N-HN1.



Fonte: Autor

A otimização da geometria foi realizada utilizando primeiramente o método PM6 (do inglês, *Parametric Method 6 Hamiltonian*) através do programa MOPAC2012¹⁰⁵. Em seguida, a otimização foi realizada através de cálculos quânticos *ab initio* ao nível Hartree-Fock utilizando como conjunto de bases o 6-31G*, utilizando o programa Orca 2.8¹⁰⁶. O conjunto de cargas calculado para a molécula mais estável obtida na otimização, foi submetido ao ajuste do potencial eletrostático clássico da molécula pelo potencial mecânico quântico usando RESP (do inglês, *Restrained Electrostatic Potential*)^{107,108}, utilizando o programa NWChem 5.1¹⁰⁹. Finalmente, diferentes ângulos torsionais foram submetidos a cálculos quânticos *ab initio* ao nível Hartree-Fock utilizando como conjunto de bases o 6-31G*, utilizando o programa Orca 2.8, e as barreiras rotacionais dos diedros foram obtidas.

Tabela 1 - Parâmetros de estiramento de ligação, flexão angular e diedral empregados no campo de força OPLS-AA.

Ligações	Tipo de Ligação	k_b (106 kJ.mol ⁻¹ .nm ⁻⁴)	b_o (nm)
(HN1,HN2)-N	H3-NT	363171.2	0.1010
C1-H11(2),C2-H21(2)	CT-HC	284512.0	0.1090
HO3-O3	HO-OH	462750.4	0.0945
(C1,C2)-N	CT-NT	319657.6	0.1448
(O1,O2)-P	O2-P	439320.0	0.1480
O3-P	OH-P	192464.0	0.1610
C1-P	CT-P	177401.6	0.1843
C3-C2	C_3-CT	265265.6	0.1522
C3-(O4,O5)	C_3-O2	548940.8	0.1250

Ângulos	Tipo de ângulo	K_θ (KJ.mol ⁻¹)	θ_o (graus)
(C1,C2)-N-(HN1,HN2)	CT-NT-H3	292.880	109.5
HN1-N-HN2	H3-NT-H3	292.880	109.5
N-C1-P	NT-CT-P	359.824	109.5
(O1,O2)-P-O3	O2-P-OH	376.560	108.23
HO3-O3-P	HO-OH-P	395.678	108.5
O3-P-C1	OH-P-CT	376.560	109.5
O1-P-O2	O2-P-O2	1171.520	119.9
C1-P-(O1,O2)	CT-P-O2	376.560	109.5
C2-C3-(O4,O5)	CT-C_3-O2	585.760	117.0
O4-C3-O5	O2-C_3-O2	669.440	126.0
C3-C2-(H21,H12)	C_3-CT-HC	292.880	109.5
C3-C2-N	C_3-CT-NT	669.440	111.2
C1-N-C2	CT-NT-CT	433.462	107.2
(H11,H12)-C1-N	HC-CT-NT	109.500	292.88
(H21,H22)-C2-N	HC-CT-NT	109.500	292.88
(H11,H12)-C1-P	HC-CT-P	109.500	343.088

Diedros	Tipo de diedro	Coeficiente de Ryckaert-Bellemans						
C3-C2-N-(HN1,HN2)	C_3-CT-NT-H3	0.72592	2.17777	0	-2.90370	0	0	0
N-C1-P-(O1,O2)	NT-CT-P-O2	0	0	0	0	0	0	0
N-C1-P-O3	NT-CT-P-OH	4.70700	-4.70700	0	0	0	0	0
(HN1,HN2)-N-C1-(H11,H12)	H3-NT-CT-HC	0.83680	2.51040	0	-3.34720	0	0	0
(HN1,HN2)-N-C2-(H21,H22)	H3-NT-CT-HC	0.83680	2.51040	0	-3.34720	0	0	0

(H11,H12)-C1-P-O3	HC-CT-P-OH	0.52300	1.56900	0	-2.09200	0	0
C1-O1-P-O2	CT-O2-P-O2	1.17570	3.52711	0	-4.70281	0	0
N-C2-C3-(O4,O5)	NT-CT-C_3-O2	3.43088	0	-	0	0	0
3.43088							
P-C1-N-C2	P-CT-NT-CT	1.78866	3.49154	0.53555	-5.81576	0	0

Para a interação dos diedros próprios foi utilizada o potencial de Ryckaert-Bellemans que descreve a função com base na expansão das potências do $\cos\varphi$ como descrito na equação (67):¹⁰⁰

$$V_{rb}(\varphi_{ijkl}) = \sum_{n=0}^5 C_n (\cos(\psi))^n \quad (67)$$

onde $\psi = \varphi - 180^\circ$.

O potencial de Ryckaert-Bellemans é equivalente para a descrição com a série de Fourier (somente as constantes são diferentes), uma vez que a série de Fourier é:

$$V_{\varphi ijkl} = F_0 + \frac{1}{2} [F_1(1 + \cos(\varphi)) + F_2(1 - \cos(2\varphi)) + F_3(1 + \cos(3\varphi)) + F_4(1 - \cos(4\varphi)) + F_5(1 + \cos(5\varphi))] \quad (68)$$

Calcula-se as constantes de Ryckaert-Bellemans através das seguintes relações:

$$\begin{aligned} C_0 &= F_0 + F_2 + \frac{1}{2}(F_1 + F_3 + F_5) \\ C_1 &= \frac{1}{2}(3F_3 - F_1 - F_5) \\ C_2 &= 4F_4 - F_2 \\ C_3 &= 10F_5 - 2F_3 \\ C_4 &= -4F_4 \\ C_5 &= -8F_5 \end{aligned} \quad (69)$$

sendo os valores encontrados adicionados ao campo de força OPLS-AA.

Desta forma, a validação do campo de força OPLS-AA usando simulação de dinâmica molecular foi realizada utilizando o programa GROMACS 4.6¹⁰². As interações eletrostáticas foram calculadas utilizando PME. A simulação foi realizada no *ensemble* NPT usando reescalonamento das velocidades com o termo estocástico para a temperatura (v-rescale) com $T=298K$ e o barostato de Berendsen a 1 bar em água explícita, com o modelo SPC por 50 ns. Todos os comprimentos de ligação foram restringidos usando o algoritmo LINCS.

As curvas dos perfis de energia calculados através das constantes obtidas para os diedros, foram obtidas através de um programa em linguagem C++ desenvolvido pelo próprio

grupo, no qual utilizou como ferramenta o método de mínimos quadrados. Sendo estas curvas comparadas aos perfis de energia obtidos pelos cálculos quânticos, para a pretendida validação.

4.3 Simulação do sistema EPSPs-GPJ-S3P

O sistema envolvendo a enzima EPSPs com seu substrato S3P, interagindo com o herbicida GPJ parametrizado no campo de força OPLS-AA, foi simulado por dinâmica molecular utilizando o programa GROMACS 4.6-beta2 em solvente explícito. A minimização de energia, bem como, a dinâmica molecular do sistema, foram realizados utilizando os mesmos parâmetros avaliados para a simulação sucedida com a enzima EPSPs em água.

4.4 Cálculo do Potencial Eletrostático e da Energia Livre de Solvatação

O potencial eletrostático calculado para a obtenção da distribuição de cargas ao longo da superfície da enzima foi obtido utilizando o programa VMD (do inglês, *visual molecular dynamics*)¹¹⁰. A energia livre de hidratação da enzima sozinha e na presença do substrato S3P foram calculadas resolvendo numericamente a equação não-linear de Poisson-Boltzmann, cuja metodologia foi implementada no programa APBS (do inglês, *Adaptive Poisson-Boltzmann Solver*)¹¹¹ em conjunto com os parâmetros ajustados do campo de força OPLS-AA. As estruturas em solução aquosa foram obtidas utilizando uma constante dielétrica de 78,54, com um raio do solvente de 1,4 nm, tensão superficial de 0,105 N.m⁻¹ e força iônica de 100 mmols.L⁻¹.

4.5 Dinâmica Molecular Direcional e Cálculo do Potencial de Força Média

A simulação computacional dos experimentos de AFM foi realizada através da Dinâmica Molecular Direcional, SMD (do inglês, *Steered Molecular Dynamics*), introduzida por Izrailev em 1997¹¹², a qual fornece modelos para explicar as forças medidas em termos das estruturas moleculares, ou seja, proporciona detalhes do processo de desvinculação induzido pela aplicação de forças externas em sistemas biomoleculares e de suas propriedades elásticas em escalas de tempo abrangidas por simulações de dinâmica molecular. A SMD tem sido largamente utilizada na investigação de funções mecânicas de proteínas tais como, processos de interação em complexos proteína-substrato¹¹³⁻¹¹⁵ e estabilidade de proteínas associadas a doenças¹¹⁶. Estas simulações darão um aporte teórico aos experimentos de AFM sob uma interpretação microscópica.

Uma forma de aplicar forças externas em um complexo proteína-ligante é restringir o

ligante em um ponto no espaço (ponto de contenção) por um potencial externo, por exemplo, o harmônico. O ponto de contenção é, então, deslocado numa direção escolhida forçando o ligante a se mover de sua posição inicial na proteína, permitindo que o mesmo explore novos contatos ao longo do caminho de desvinculação. Assumindo uma única coordenada de reação x , e um potencial externo $V=k(x-x_0-vt)^2/2$, onde k é a rigidez do ponto de contenção, e x_0 é a posição inicial do ponto de contenção movendo com uma velocidade constante, v , a força externa no sistema pode ser expressado como¹¹⁷,

$$F=k(x_0+vt-x) \quad (70)$$

força F , correspondente à molécula sendo puxada por uma mola harmônica de rigidez k com uma velocidade constante v .

Foram realizadas simulações distintas onde a molécula de glifosato foi afastada do sistema EPSPs-S3P-GPJ ao longo de três possíveis caminhos de retirada deste herbicida do sítio ativo da enzima. A metodologia utilizada foi o *constraint pulling*, disponível no programa GROMACS, na qual a distância entre os centros de massa dos dois grupos (grupo 0 → EPSPs-S3P e grupo 1 → GPJ) foi restringida com uma constante de força, k , de 367 KJ.mol⁻¹.nm⁻¹, onde em todas as dinâmicas a velocidade de afastamento, v , foi de 0,001 nm.ps⁻¹.

A dinâmica do sistema foi realizada em um *ensemble* NVT utilizando-se o algoritmo de *leap-frog*. Foi utilizado um intervalo de tempo de integração de 1 fs totalizando um tempo de simulação de 4 ns, com a temperatura de 298 K, utilizando termostato (v-rescale).

Neste estudo, cálculos de Potencial de Força Média, PMF (do inglês, *Potential of Mean Force*) foram realizados para determinar a energia livre absoluta de interação entre EPSPs-S3P e o GPJ, com o intuito de quantificar a interação entre a ponta do AFM funcionalizada com a enzima EPSPs e o glifosato.

O PMF é um potencial que é obtido pela integração da força média de um *ensemble* de configurações, no qual exibe um importante papel na investigação de processos moleculares, em que o espaço configuracional é descrito por uma coordenada de reação. Assim, o PMF é basicamente o perfil de energia livre ao longo da coordenada de reação e é determinado através da média ponderada de Boltzmann sobre todos os graus de liberdade exceto o da coordenada de reação.

O SMD é um método efetivo para explorar processos mecânicos e moleculares

acessíveis através de experimentos de AFM. No entanto, uma simulação dinâmica direcional é um processo de não-equilíbrio, enquanto que o PMF é uma propriedade em equilíbrio. Portanto, se torna necessária uma teoria que conecte os processos em equilíbrio e não-equilíbrio, que se dá através da mecânica estatística em não-equilíbrio, especialmente através da igualdade de Jarzynski¹¹⁸. Se torna possível desta forma, extrair propriedades em equilíbrio de sistemas em não-equilíbrio. A igualdade de Jarzynski estabelece uma conexão entre o cálculo de energia livre em equilíbrio e o trabalho realizado no processo em não-equilíbrio, e portanto, permite calcular o PMF sobre processos em não-equilíbrio tais como as simulações de SMD¹¹⁴.

A igualdade de Jarzynski é:

$$e^{-\beta \Delta F} = \langle e^{-\beta W} \rangle \quad (71)$$

onde ΔF representa a variação de energia livre entre os estados inicial e final, W é o trabalho externo e $\beta = (K_B T)^{-1}$ em que K_B e T são a constante de Boltzmann e a temperatura, respectivamente. A propriedade mais importante desta relação é que ela não se restringe somente a regimes em equilíbrio. A igualdade é satisfeita para qualquer perturbação, desde que amostragens suficientes sejam realizadas¹¹⁹.

Uma expansão cumulante da equação de Jarzynski pode ser utilizada para contornar a dificuldade de estimar a média da exponencial, lado direito da equação 71. O logaritmo de uma média de uma exponencial pode ser expandida em termos de cumulantes,

$$\log \langle e^x \rangle = \langle x \rangle + \frac{1}{2} (\langle x^2 \rangle - \langle x \rangle^2) + \dots \quad (72)$$

onde são mostrados o primeiro e segundo cumulantes na equação (72). Nesse caso, em que todos os outros termos cumulantes são negligenciados, a variável x deve ser, impreterivelmente, amostrada através de uma distribuição gaussiana. Usando esta expansão, a energia livre pode ser obtida através da equação 73,

$$F_{\lambda(\tau)} - F_{\lambda(0)} = \langle W(\tau) \rangle + \frac{\beta}{2} (\langle W(\tau)^2 \rangle - \langle W(\tau) \rangle^2) + \dots \quad (73)$$

onde, $F_{\lambda(\tau)} - F_{\lambda(0)}$ é a variação de energia livre, $\beta = 1/k_B T$ e $W(\tau)$ é o trabalho. Quando o

trabalho segue uma distribuição gaussiana, a fórmula até segunda ordem pode ser utilizada, sendo os cumulantes de terceira e maiores ordens iguais a zero^{120,121}.

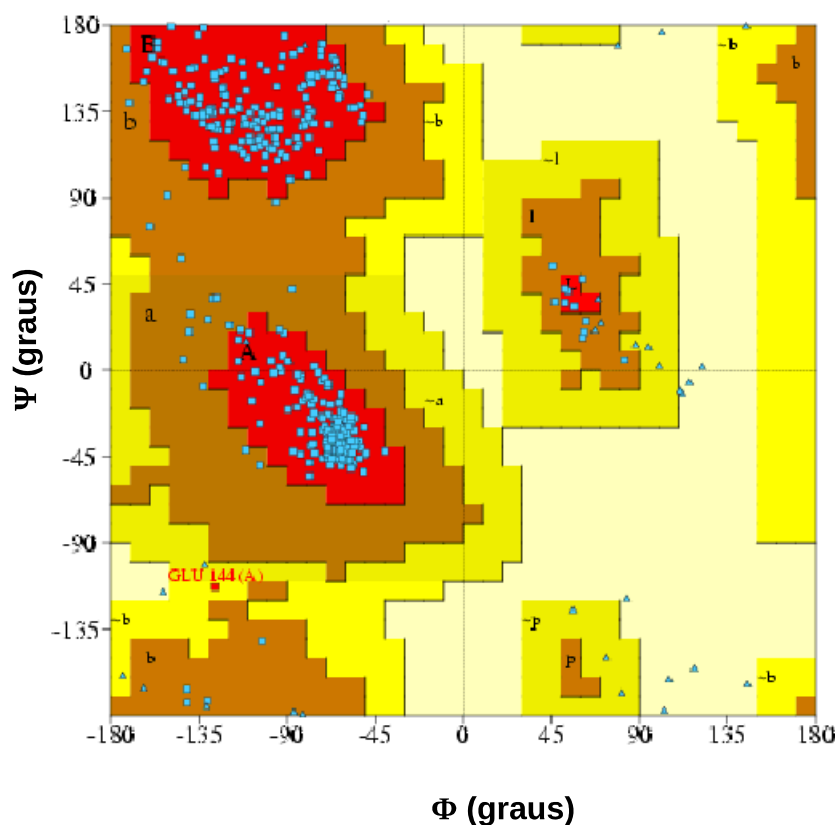
Após determinado o(s) caminho(s) mais favorável(veis), foram realizadas 80 simulações SMD, com velocidade de *pulling* de 0,001 nm.ps⁻¹ e com uma constante de força de 367 KJ.mol⁻¹.nm⁻¹. O trabalho bem como o potencial de força média foram obtidos, para cada um dos diferentes eixos escolhidos, utilizando o software Octave 3.8.1¹²². Foi realizado o teste de Shapiro-Wilk¹²³ para avaliar se a distribuição do conjunto de dados, referente aos valores do trabalho obtidos nas simulações ao longo de cada eixo, adere à distribuição Normal, utilizando o software R-Statistics 3.0.2¹²⁴.

5 Resultados e Discussões

5.1 Caracterização da dinâmica estrutural e do potencial eletrostático da enzima EPSPs em solução aquosa.

A qualidade estereoquímica do modelo 3D de uma proteína pode ser indicada pela distribuição dos ângulos torcionais Φ e ψ da cadeia principal. A qualidade do modelo está ligada à escolha do molde e à construção do alinhamento, que servirá de referência para a construção dos átomos que compõem a proteína. A Figura 23, mostra a distribuição de todos estes ângulos na enzima EPSPs através do diagrama de Ramachandran, obtidas através do PBDsum¹²⁵. A análise do diagrama de Ramachandran indica claramente que a estrutura de coordenadas atômicas apresenta uma ótima estatística estereoquímica com 90,8% dos resíduos nas regiões favoráveis (em vermelho), 8,9% nas regiões adicionalmente permitidas (em marrom), 0,3% nas regiões generosamente permitidas (em amarelo) e 0,0% nas regiões desfavoráveis (em amarelo-claro), Tabela 2.

Figura 23 - Diagrama de Ramachandran indicando as distribuições dos ângulos torcionais Φ e ψ da cadeia principal. As regiões mais favoráveis são mostradas em vermelho, as adicionalmente permitidas em marrom, as generosamente permitidas em amarelo e as não permitidas em amarelo claro.



Fonte: ¹²⁵

Os resíduos de glicina (totalizando 37 resíduos) devem ser desconsiderados (apresentados na Figura 23 como triângulos), uma vez que não possuem cadeia lateral, os de prolina (totalizando 18 resíduos), pois nestes o carbono alfa está ligado à cadeia lateral e os resíduos das extremidades, C-terminal e N-terminal (totalizando 2 resíduos), que apresentam padrões estereoquímicos diferentes dos outros resíduos.

Tabela 2 - Estatística da estereoquímica do modelo 3D da enzima EPSPs.

	Número de resíduos	Porcentagem (%)
Resíduos nas regiões mais favorecidas (A, B, L)	336	90,8
Resíduos nas regiões adicionalmente permitidas (a, b, l, p)	33	8,9
Resíduos nas regiões generosamente permitidas (-a, -b, -l, -p)	1	0,3
Resíduos nas regiões proibidas	0	0
	370	100

Fonte: ¹²⁵

A análise do gráfico de Ramachandran, por si só, não valida a estrutura de uma proteína, mas é uma ferramenta valiosa na análise da estrutura 3D, visto que, abre as portas para os estudos de base molecular entre esta proteína e seu respectivo ligante.

Após a Dinâmica Molecular, a análise da estabilidade conformacional da enzima EPSPs (Figura 24) foi avaliada através da medida da raiz quadrada do desvio quadrático médio (RMSD, do inglês, *Root Mean Square Deviation*), o qual revela o distanciamento conformacional da estrutura nativa ao longo do tempo de simulação e também pela medida do seu raio de giro (RG), onde pode-se obter uma medida aproximada da compactação da estrutura da enzima.

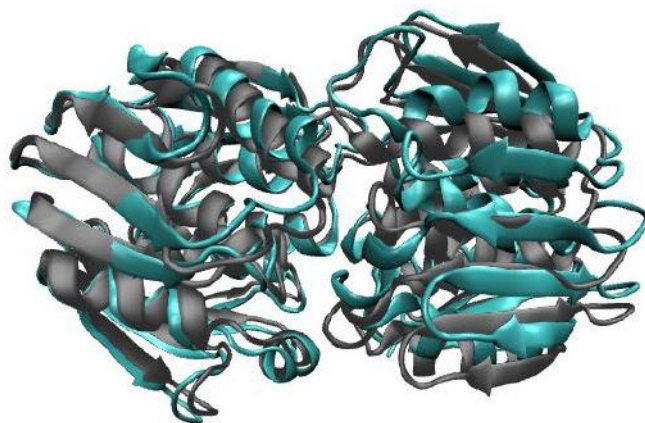
Para o cálculo dos desvios entre a conformação da enzima ao decorrer da dinâmica molecular em relação a uma estrutura de referência, ajusta-se pelos mínimos quadrados a estrutura em relação à referência ($t_2=0$), pela equação,

$$RMSD = \left[\frac{1}{M} \sum_{i=1}^N m_i \|r_i(t_1) - r_i(t_2)\|^2 \right]^{\frac{1}{2}} \quad (74)$$

onde $M = \sum_{i=1}^N m_i$ e $r_i(t)$ é a posição do átomo i no tempo t e N é o número de átomos

utilizado no cálculo.

Figura 24 - Representação da sobreposição da estrutura da enzima EPSPs (representação new cartoon) , antes da simulação por dinâmica molecular (em cinza) e após a simulação (em verde).



Para se observar a compactação da estrutura da enzima, utiliza-se o raio de giro (RG), dado pela equação,

$$R_g = \left(\frac{\sum_{i=1}^N \|r_i\|^2 m_i}{\sum_{i=1}^N m_i} \right)^{\frac{1}{2}} \quad (75)$$

onde m_i é a massa do átomo i e r_i é sua posição em relação ao centro de massa.

A enzima EPSPs apresentou pequena mobilidade estrutural em solução aquosa (Figura 25), à 298 K, manifestando uma leve e sistemática flutuação tanto para as medidas de RMSD, em torno de 0,25 nm ($\pm 0,015$ nm) (Figura 25-A), quanto para as medidas de RG, em torno de 2,16 nm ($\pm 6,8 \cdot 10^{-3}$ nm) (Figura 25-B), valores os quais não representam uma alteração significativa na conformação da enzima inicial. Isto sugere que a proteína provavelmente não perderá sua atividade durante a deposição da mesma no cantiléver do AFM ou durante a detecção do herbicida no nanobiossensor proposto.

A análise do potencial eletrostático da enzima EPSPs revela uma uniformidade na distribuição de cargas na sua superfície, de maneira geral (Figura 26). No entanto, no sítio ativo, ocorre uma maior distribuição de cargas negativas na região ao longo do eixo y, voltado para frente, (Figura 26-B). Uma vez identificado o possível caminho de retirada do herbicida glifosato do sítio ativo (que será realizado na seção 5.4), pode-se sugerir uma funcionalização da ponta do AFM de maneira que grupos funcionais carregados positivamente ou negativamente possam ser inseridos na ponta do cantiléver do equipamento.

Figura 25 - (A) Raiz quadrada do desvio quadrático médio - RMSD e (B) raio de giro - RG da enzima EPSPs.

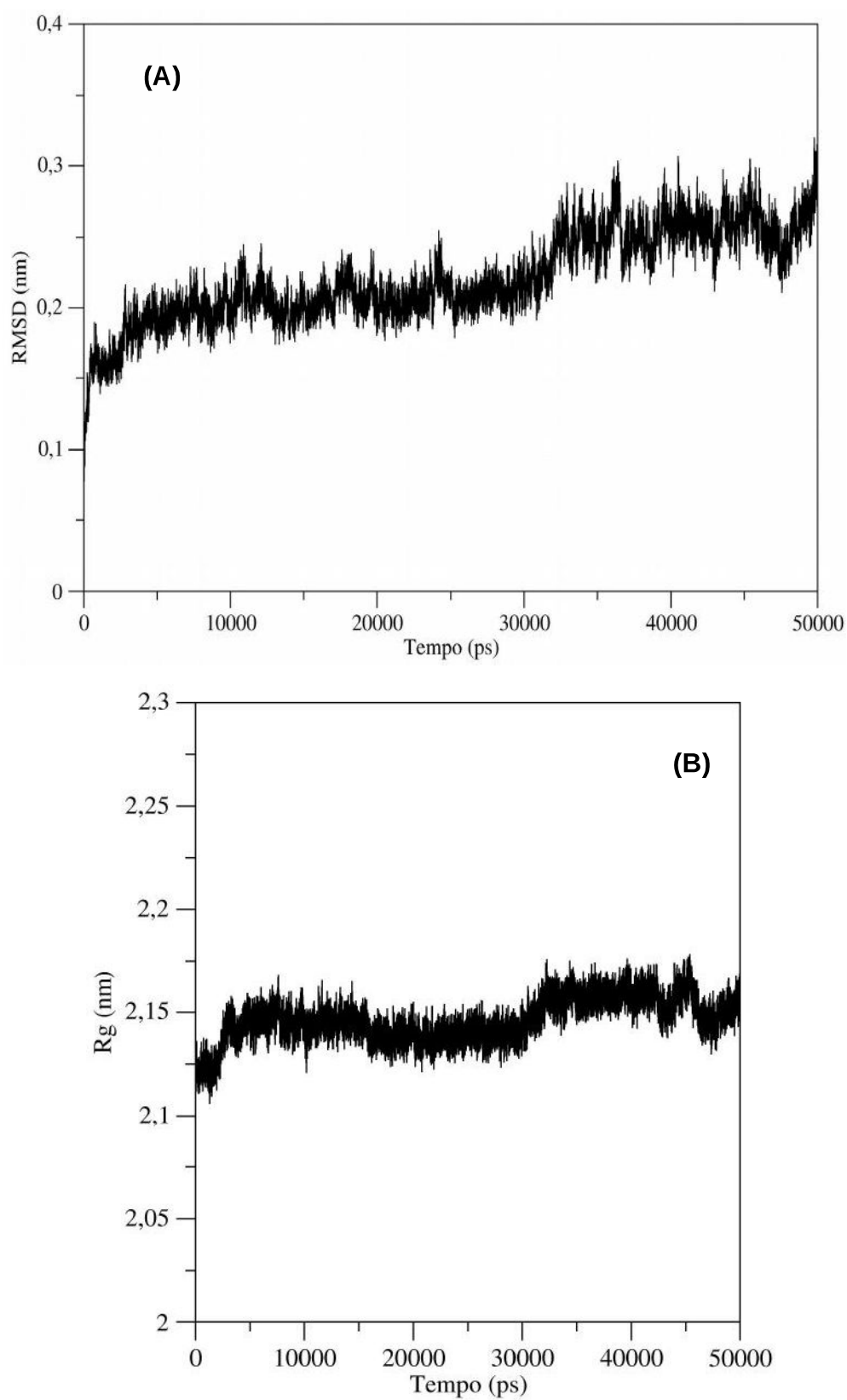
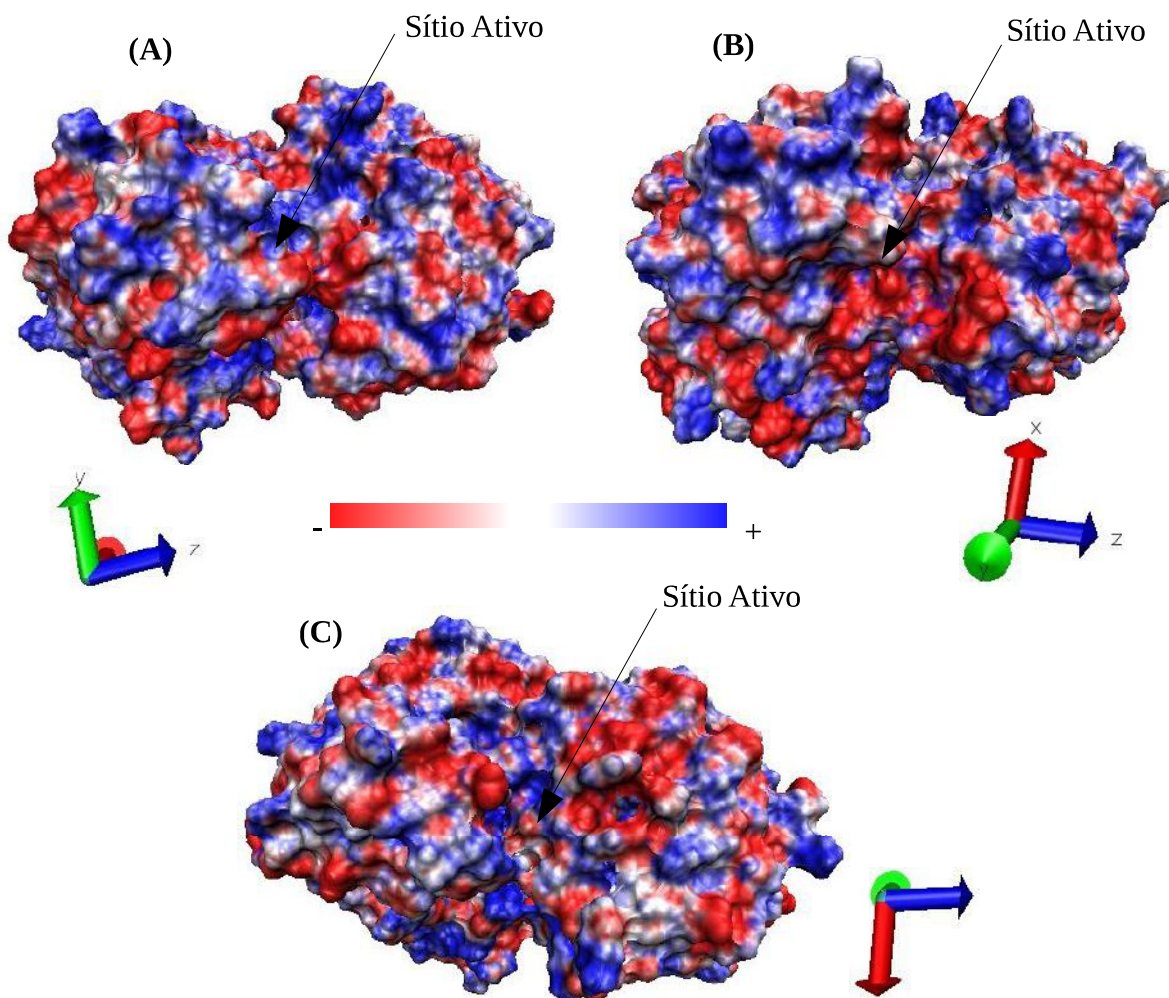


Figura 26 - Potencial eletrostático plotado na superfície molecular da enzima EPSPs. Superfície no eixo -x voltado para frente (A); superfície no eixo y voltado para frente (B) e superfície no eixo -y voltado para frente (C).



5.2 Desenvolvimento de um modelo molecular para o herbicida e parametrização de campos de força

Nas simulações por dinâmica molecular, os parâmetros do campo de força do sistema devem ser definidos para reproduzir dados energéticos e estruturais. Portanto, há a necessidade de se desenvolver um modelo que reproduza as medidas experimentais com maior grau possível de precisão. Em geral, as parametrizações incluem interações entre átomos ligados e átomos não-ligados. Neste trabalho, utilizou-se cálculos quânticos para obter as cargas atômicas parciais e as constantes de força que reproduzem as curvas de energia potencial de dois diedros da molécula de glifosato, que foram parametrizados no campo de força OPLS-AA.

Cálculo das cargas:

Após a otimização da estrutura do glifosato, as cargas parciais da molécula foram ajustadas para o potencial eletrostático no campo de força OPLS-AA usando o método RESP e os valores estão apresentados na Tabela 3.

Tabela 3 - Carga parcial dos átomos do glifosato.

Átomos	Tipo	Carga (e*)
P1	PO ₃ ⁻	1.342
O1	O	-0.860
O2	O	-0.860
O3	OH	-0.753
HO3	H	0.417
C1	CH ₂	-0.164
H11	H	0.066
H12	H	0.066
N1	N ⁺	0.110
HN1	H	0.178
HN2	H	0.178
C2	CH ₂	-0.103
H21	H	0.045
H22	H	0.045
C3	CO ⁻	0.879
O4	O	-0.793
O5	O	-0.793
Carga Total		-1,000

*e = 1,602.10⁻¹⁹ C

Dinâmica molecular:

A dinâmica molecular do herbicida em solução aquosa foi realizada para validar os parâmetros dos diedros HO3-O3-P-C1 e P-C1-N-HN1 no campo de força OPLS-AA obtidos neste estudo. Os coeficientes e os ângulos dos diedros foram calculados comparando uma varredura mecânico-quântica da energia conformacional com uma estrutura equivalente para cálculos utilizando mecânica molecular. A estrutura (Figura 27) mostrou-se estável durante o tempo de simulação com valores de RMSD de 0.01-0.15 nm para a molécula de glifosato, (Figura 28).

Figura 27 - Representação da sobreposição da molécula glifosato (representada em CPK) antes e após a simulação por dinâmica molecular.

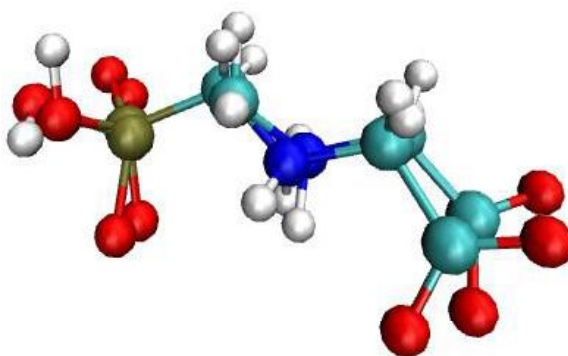
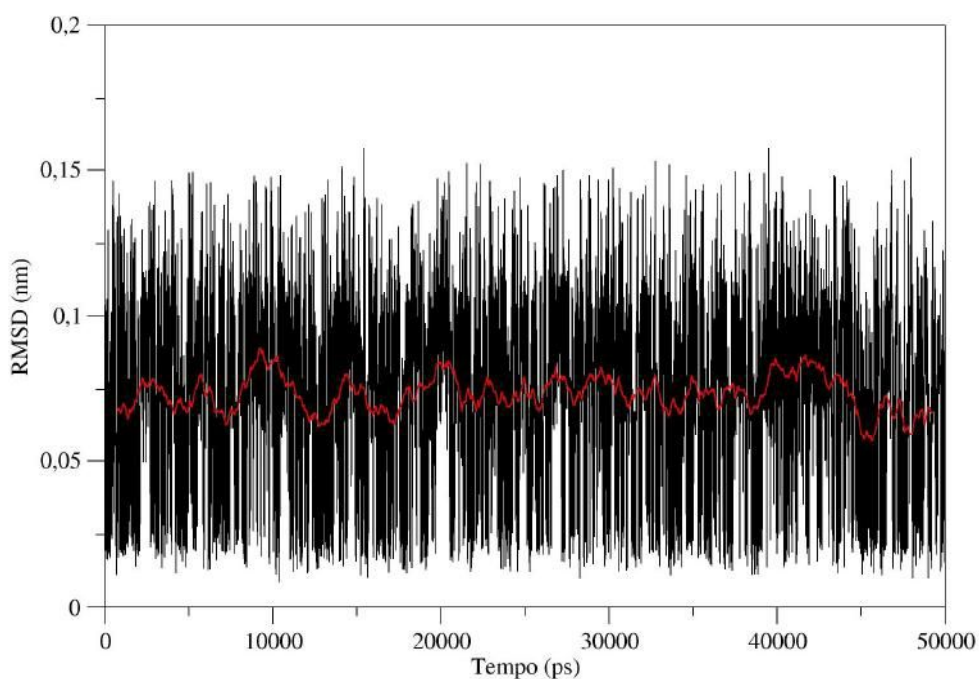


Figura 28 - Raiz quadrada do desvio quadrático médio (RMSD) do herbicida glifosato.



A Tabela 4 apresenta os valores dos coeficientes e dos ângulos dos mínimos de energia dos diedros calculados com QM e MM.

A Figura 29 apresenta as superfícies de energia livre dos diedros do glifosato obtidos através dos cálculos quânticos e dos cálculos clássicos utilizando as constantes obtidas.

Tabela 4 - Coeficientes e ângulos dos diedros parametrizados.

Diedros	Coeficientes no campo OPLS-AA						
Parametrizados							
HO3-O3-P-C1	7.19185	20.5797	-15.0898	1.27049	-0.984426	0.269785	
P-C1-N-HN1	19.3728	-17.6137	7.4594	3.17068	1.29439	1.88262	
Diedros	Dinâmica	Cálculo		HF/6-31G*			
Parametrizados	Molecular	clássico***		Ψ(graus)			
	Ψ(graus)	Ψ(graus)					
HO3-O3-P-C1	101/256	109/239		119/259			
P-C1-N-HN1	208/331	240/340		210/330			

*** Fórmula utilizada:

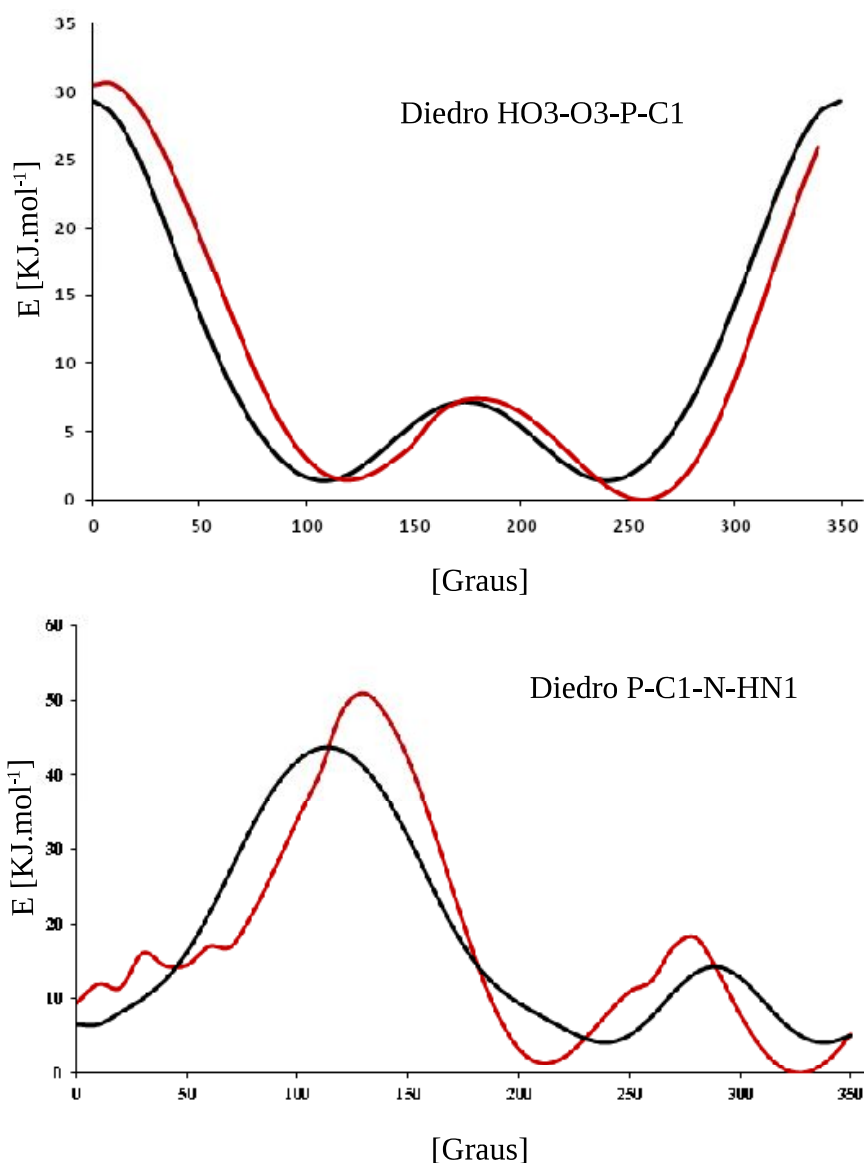
$$F(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k) = \sum_{k=1}^m [f(x_k) - \Psi(x_k)]^2$$

$$\Psi(x_k) = \alpha_1 h_1(x_k) + \alpha_2 h_2(x_k) + \alpha_3 h_3(x_k) + \dots + \alpha_m h_m(x_k)$$

Para o diedro HO3-O3-P-C1 ambos os cálculos apresentaram dois mínimos de energia, sendo que a menor energia se dá para o ângulo de 259° no cálculo químico-quântico e 239° no cálculo clássico e o segundo mínimo de energia ocorre em 119° com QM e 109° com o clássico. Ambos os mínimos de energia ocorreram em ângulos menores nos cálculos clássicos, sendo que o menor mínimo de energia apresenta uma diferença de energia de aproximadamente 1.1 KJ.mol⁻¹, com relação ao cálculo quântico. O potencial de Ryckaert-Bellemans para a molécula foi de 49,98 KJ.mol⁻¹ com desvio padrão de 5,32 KJ.mol⁻¹, portanto, a diferença entre os mínimos se encontra dentro do desvio.

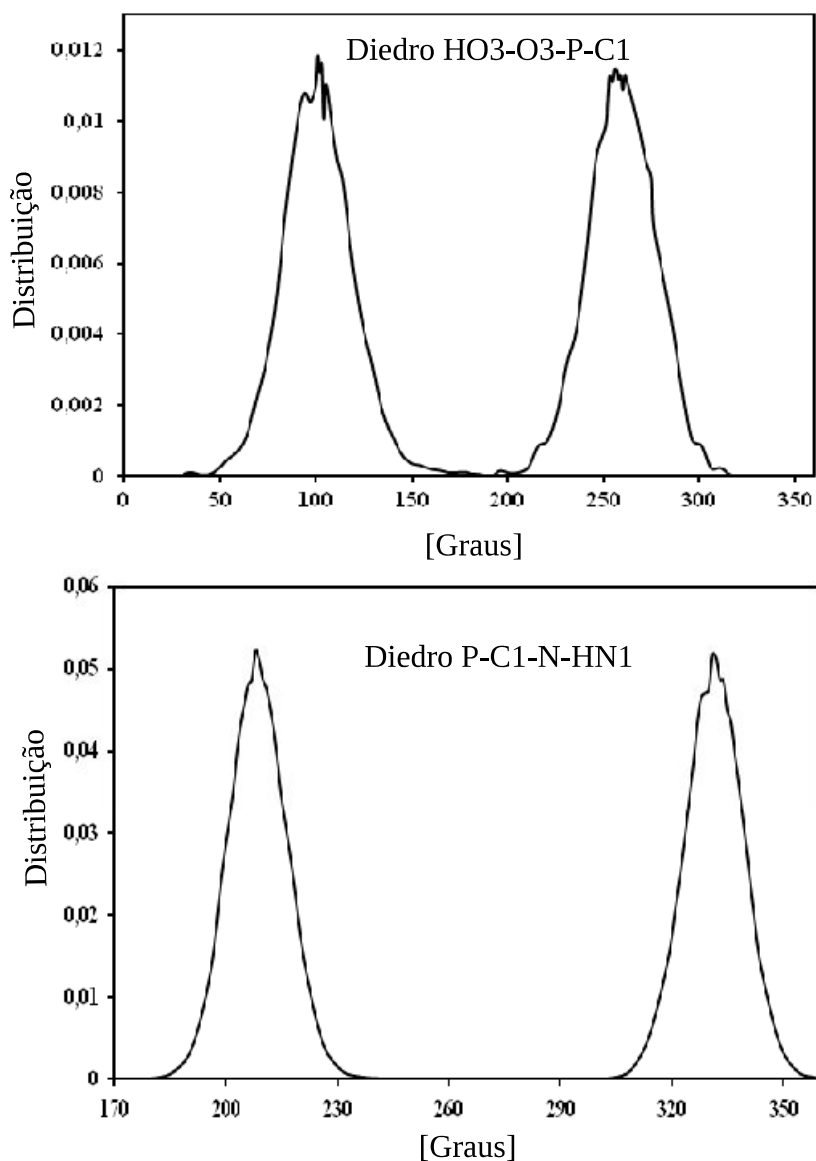
Para o diedro P-C1-N-HN1 ambos os cálculos apresentaram dois mínimos de energia que estão relacionados com os dois diedros formados com os diferentes hidrogênios (P-C1-N-HN1 e P-C1-N-HN2) em que a menor energia se deu com o ângulo de 330° com QM e 340° com método clássico com uma diferença de energia em torno de 4.2 KJ.mol⁻¹ (dentro do desvio padrão) e o segundo mínimo em 210° com QM e 240° com o método clássico apresentando também uma proximidade entre os métodos.

Figura 29 - Perfil de energia calculados com cálculo quântico (vermelho) e cálculo clássico usando o método dos mínimos quadrados (preto) para os diedros HO3-O3-P-C1 e P-C1-N-HN1.



A Figura 30 apresenta a distribuição de probabilidade dos ângulos após a dinâmica molecular em solução aquosa. No diedro HO3-O3-P-C1 ocorreu duas distribuições principais, sendo uma no ângulo de 256° e outra com ângulo de 101°, onde o mínimo de energia nos cálculos quânticos ocorreu no ângulo de 259°, coincidindo com uma das distribuições. No diedro P-C1-N-HN1 também ocorreram duas distribuições principais, nos ângulos 331° e 208°, ângulos muito próximos dos mínimos de energia obtidos nos cálculos quânticos. Desta forma, estes parâmetros foram utilizados para as simulações envolvendo o herbicida glifosato na enzima EPSPs.

Figura 30 - Distribuição do ângulo diedral, para o diedro HO3-O3-P-C1 (acima) e para o diedro P-C-N-HN1 (abaixo).



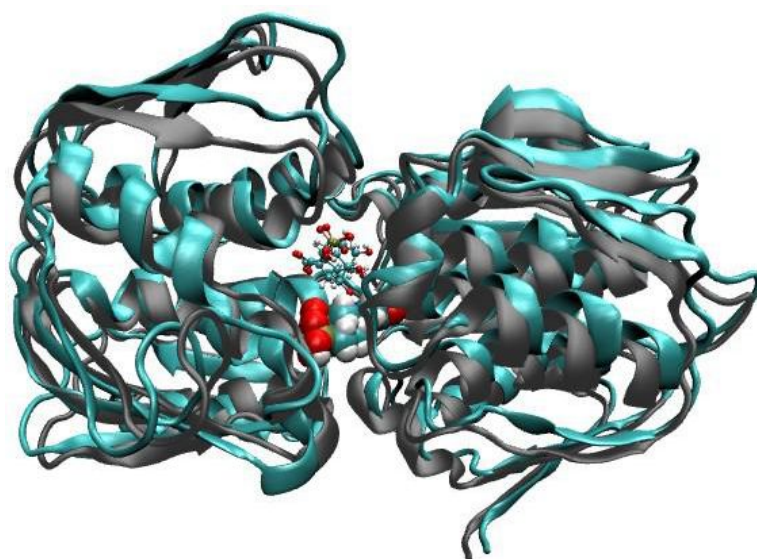
5.3 Caracterização do sistema EPSPs-S3P-GPJ

A dinâmica molecular do sistema contendo a enzima EPSPs, o substrato shikimato-3-fosfato e o herbicida glifosato (Figura 31) foi realizada e algumas propriedades foram calculadas.

O biossensor pretendido requer uma imobilização da enzima EPSPs na ponta do cantiléver do AFM para detectar a interação com o herbicida glifosato. De acordo com a literatura a EPSPs pode apresentar dois tipos de conformação, aberta e fechada^{9,10}. Segundo Schönbrunn e colaboradores¹⁰, o substrato S3P é o responsável pela transição do estado aberto para o fechado. Desta forma, é necessário conhecer as características intrínsecas da estrutura

da enzima em solução para selecionar a estrutura enzimática adequada para interagir com o glifosato. Esta interação é fortemente relacionada com a flutuação conformacional da enzima em solução, sendo esta dinâmica estrutural avaliada através do cálculo da RMSD.

Figura 31 - Representação da sobreposição do sistema EPSPs-S3P-GPJ. Enzima (representada em *newcartoon*), shikimato-3-fosfato (em CPK) e o herbicida glifosato (representado em superfície de van der Waals). EPSPs antes da simulação por dinâmica molecular (em cinza) e após a dinâmica molecular (em verde).

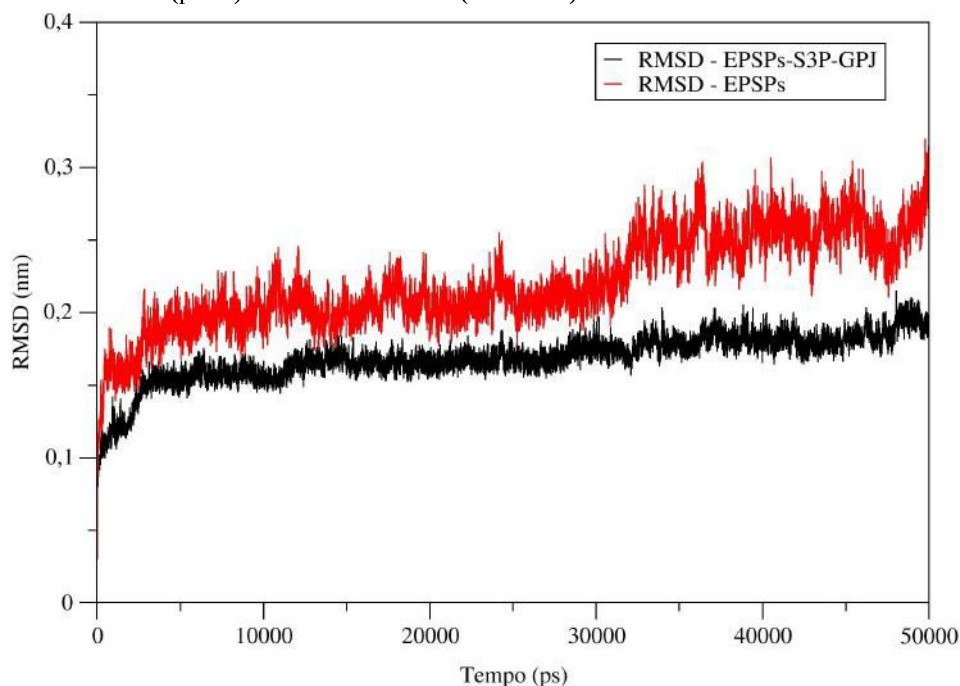


A análise do RMSD da enzima no complexo mostra uma pequena mobilidade estrutural, em torno de 0,17 nm ($\pm 0,011$ nm). A flutuação conformacional da enzima na presença do substrato e do herbicida foi menor do que na enzima isolada (Figura 32), fato que propõe uma estabilização ocorrida devido à presença do substrato S3P, como sugerido anteriormente.

Tanto na dinâmica molecular da enzima isolada quanto da enzima no complexo, em solução aquosa, a conformação estrutural se mostrou estável. No entanto, pode-se deduzir que a força de interação entre a estrutura na presença do substrato S3P no sítio ativo e o herbicida glifosato é mais provável de ser transmitida para a ponta do AFM devido à maior rigidez desta estrutura, evitando assim, a maior parte das perturbações mecânicas produzidas pelas interações entre a enzima e o herbicida.

Para descrever os efeitos da solvatação na flutuação estrutural e na estabilização da estrutura da enzima EPSPs (isolada e na presença do substrato), a energia livre de solvatação foi calculada via resolução numérica da equação não-linear de Poisson-Boltzman.

Figura 32 - Raiz quadrada do desvio quadrático médio (RMSD) da EPSPs na presença do substrato S3P e do inibidor GPJ (preto) e da EPSPs isolada (vermelho).



A energia livre de solvatação para o sistema enzima-substrato é mais favorável quando comparada à energia livre para a enzima sozinha (Tabela 5), ou seja, a água atuou favoravelmente para a formação do complexo EPSPs-S3P de forma mais significativa do que no ambiente na presença apenas da enzima.

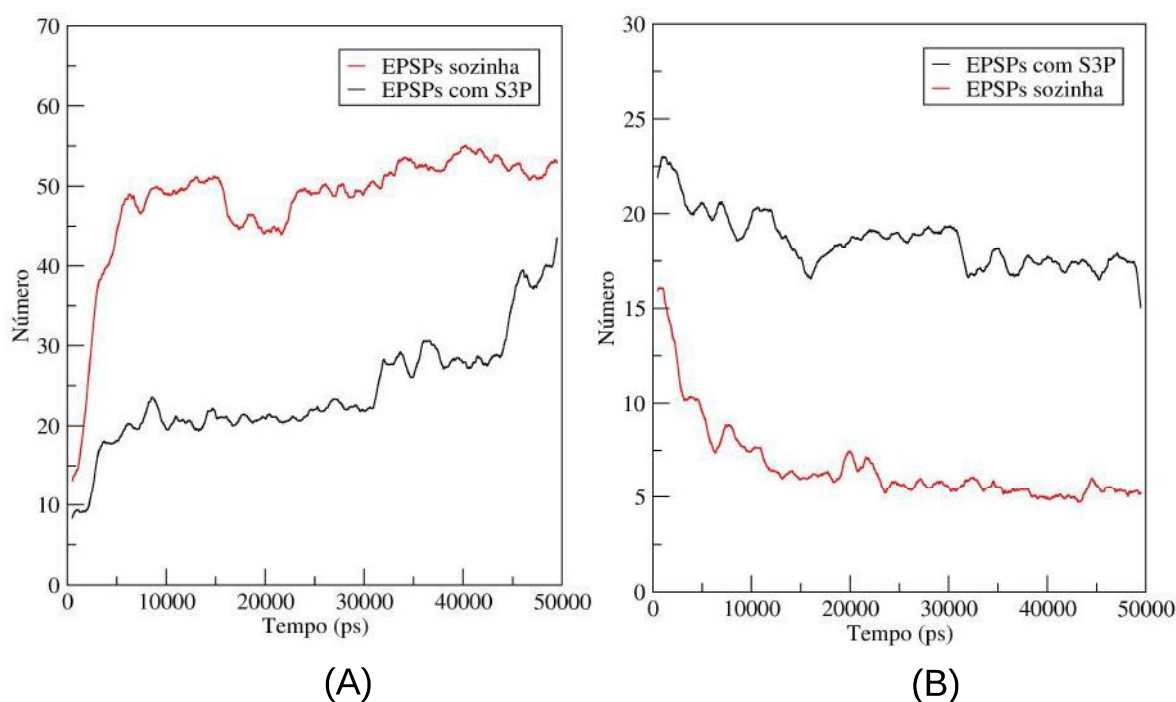
Tabela 5 - Energia livre de solvatação da enzima EPSPs com o substrato S3P e da enzima sozinha.

Sistema	Sistema (EPSPs-S3P)	EPSPs sozinha
ΔG solvatação (KJ.mol ⁻¹)	-1,402x10 ⁴	-1,326x10 ⁴

Na Figura 33-A, observa-se que o número de ligações de hidrogênio entre o sítio ativo da enzima sozinha e a água é maior quando comparado com a enzima na presença do substrato. Além disso, as interações são formadas rapidamente após o início da dinâmica nesse primeiro sistema, o que sugere uma maior abertura do sítio ativo sem a presença do substrato, ou seja, uma estrutura da enzima na forma mais aberta. Na presença do S3P a formação das ligações de hidrogênio ocorrem de forma mais gradual, sugerindo uma estrutura mais fechada, estabilizada pelo substrato. A Figura 33-B, apresenta as interações de hidrogênio intermoleculares entre os resíduos do sítio ativo. O número de interações na presença do substrato é acentuadamente maior do que quando o mesmo está ausente. A entrada de água no sítio ativo da EPSPs na presença do S3P, bem como o rompimento das

interações intermoleculares entre os resíduos é dificultada devido à conformação mais fechada da enzima.

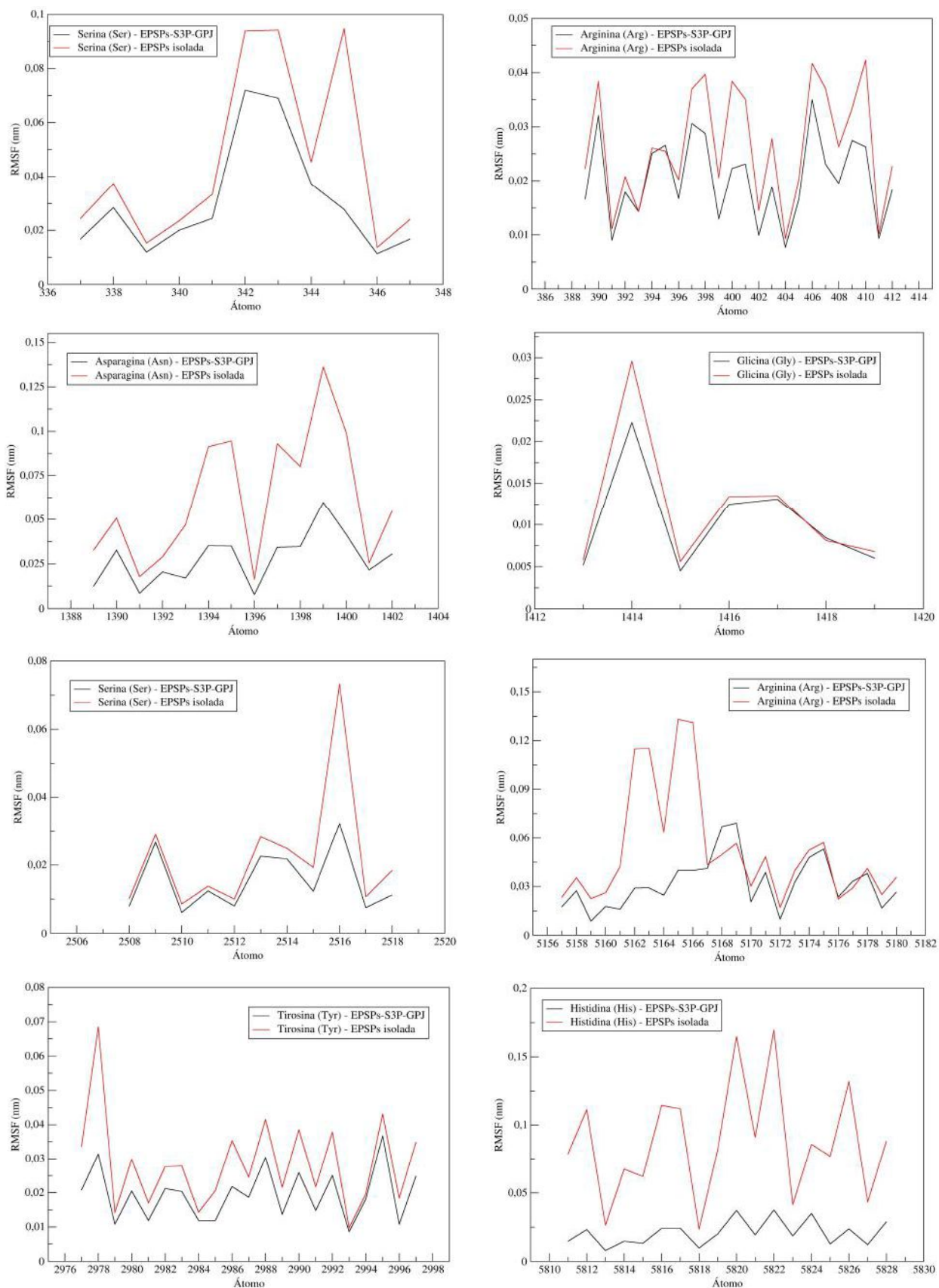
Figura 33 - (A) Número de ligações de hidrogênio entre a enzima sozinha (vermelho) e a água e entre a enzima com S3P e água (preto) e (B) interações intermoleculares entre os resíduos do sítio ativo na enzima sozinha (vermelho) e na enzima com S3P (preto).



A raiz quadrada da flutuação quadrática média RMSF (do inglês, *Root Mean Square Fluctuation*) dos resíduos de aminoácidos do sítio ativo da EPSPs foi avaliada com o intuito de verificar a variação da posição dos átomos de cada resíduo durante a dinâmica. Dos 19 resíduos de aminoácidos presentes no sítio ativo da enzima, 11 resíduos apresentaram diferenças desprezíveis entre o sistema com a EPSPs isolada e o sistema EPSPs-S3P-GPJ. A Figura 34, mostra a variação da posição dos átomos dos 8 resíduos de aminoácidos que compõem a região do sítio ativo da enzima EPSPs em que, as flutuações das posições dos átomos que ocorrem no sistema EPSPs-S3P-GPJ são claramente menores quando comparadas às variações ocorridas no sistema com a enzima isolada. As interações entre o Shikimato-3-Fosfato e estes resíduos de aminoácidos restringem a movimentação dos átomos, demonstrando uma especificidade típica de um sistema enzima-substrato.

A partir das informações obtidas acima, sugere-se que a funcionalização da ponta do AFM seja composta pela enzima na presença do seu substrato S3P, para produzir um nanobiossensor e consequentemente detectar o herbicida glifosato.

Figura 34 - Flutuação quadrática média (RMSF) de 8 resíduos de aminoácidos do sítio ativo na presença do S3P-GPJ (preto) e na enzima isolada (vermelho).

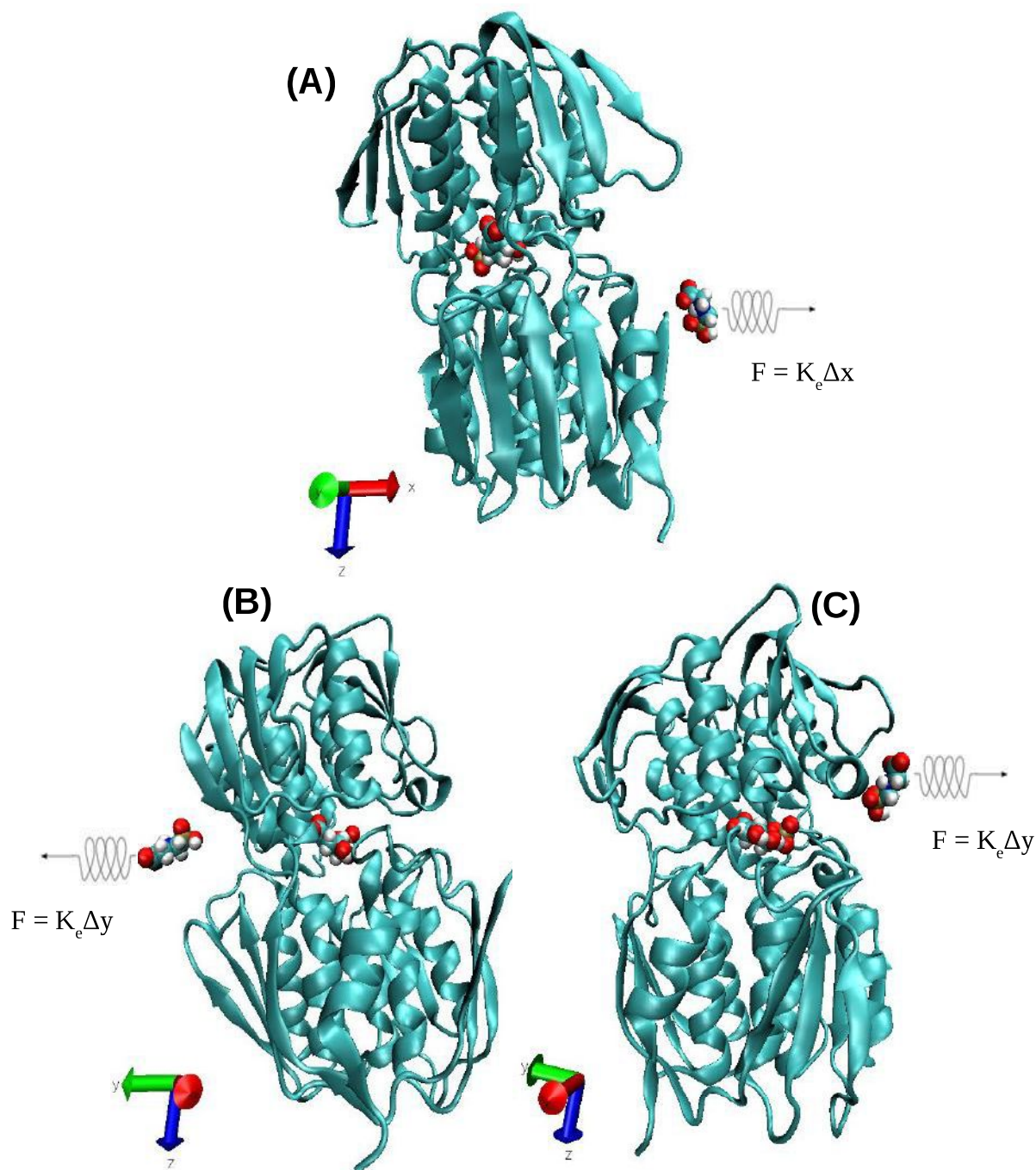


5.4 Descrição atomística e previsões moleculares no processo de detecção do herbicida pelo nanobiossensor

As simulações via Dinâmica Molecular Direcional (SMD) são orientadas pela adição de forças externas especificadas pelo usuário. O propósito da introdução destas forças externas é ajudar o sistema molecular a superar as barreiras de energia entre estados.

As SMD requerem portanto, a seleção de um caminho, isto é, uma série de direções ou uma direção única em que as forças serão aplicadas. Para determinar um possível caminho de desvinculação do glifosato do sítio ativo, a dinâmica molecular direcional foi realizada de forma que o herbicida fosse puxado para fora da enzima a partir de sua posição inicial no sítio ativo ao longo dos eixos x , y e $-y$ (Figura 35). Essas direções foram escolhidas devido ao menor impedimento estérico.

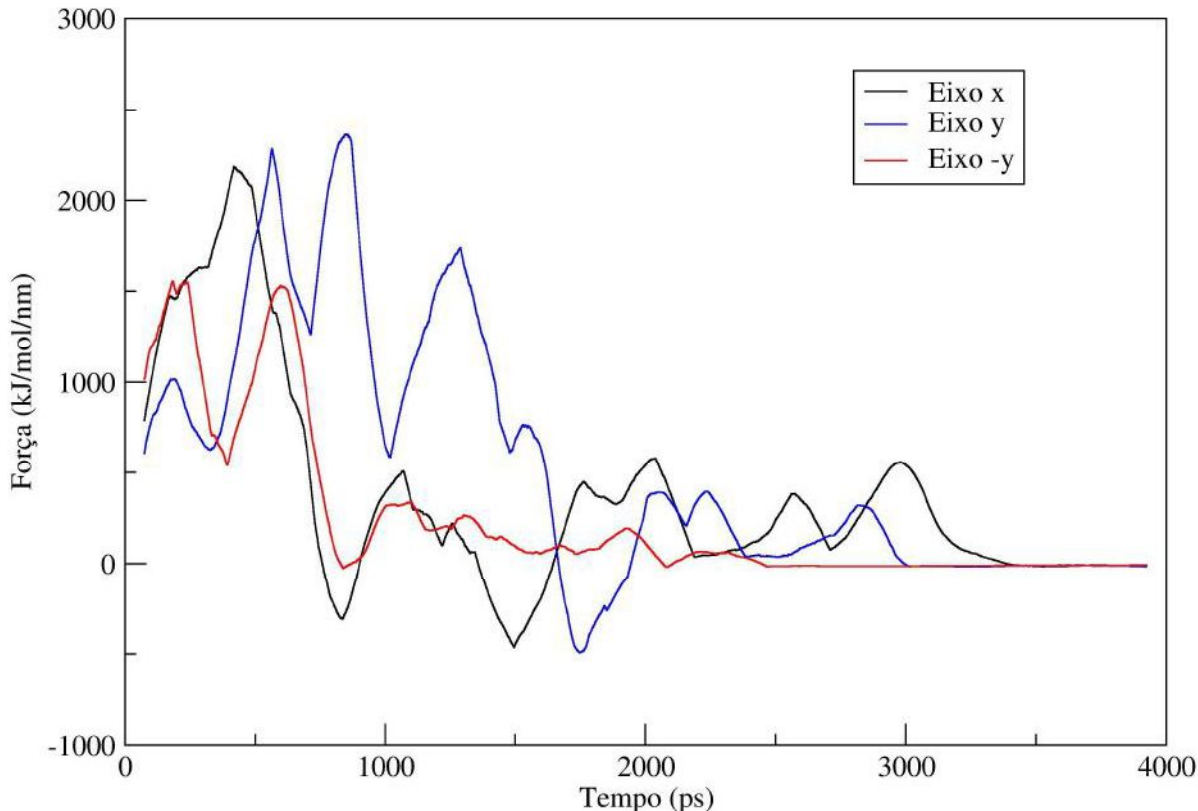
Figura 35 - Dinâmica Molecular Direcional em três possíveis caminhos de retirada do herbicida glifosato. (A) Ao longo do eixo x; (B) Ao longo do eixo y e (C) ao longo do eixo -y.



A partir da análise do gráfico de força em função do tempo (Figura 36), os caminhos ao longo dos eixos x , y e $-y$ apresentaram forças máximas nos valores aproximados de 2150, 2350 e 1530 $\text{kJ mol}^{-1} \text{nm}^{-1}$ (3570, 3900 e 2540 pN), respectivamente. Esses valores são da

mesma ordem de magnitude de simulações de SMD realizadas em trabalho anterior do grupo, no qual a força de extração varia na faixa de 1100 a 2100 pN¹¹⁵. O caminho de retirada do herbicida que necessita de uma menor força foi ao longo do eixo -y, porém o mesmo apresentou dois momentos de força máxima, a 200 e 600 ps, que estão relacionados às coordenadas onde as interações de hidrogênio entre o glifosato e os resíduos do sítio ativo EPSPs-S3P são mais intensas. A retirada ao longo do eixo x apresentou apenas um pico de força máxima significativo, a 430 ps, a partir do qual pode ser atribuído o ponto de ruptura das principais interações do herbicida com o sítio ativo. Já o caminho ao longo do eixo y apresentou a maior força de interação, a 850 ps, entre os três caminhos, além de ter ocorrido mais outros dois picos significativos, 2250 e 1730 kJ mol⁻¹ nm⁻¹ (3740 e 2870 pN), a 575 e 1300 ps. Todos esses picos de força mencionados nos três eixos estão relacionados com as interações do glifosato ainda com resíduos do sítio ativo. Na sequência desses sinais, todos os traços de força, que são acentuadamente mais fracos, estão relacionados com as interações do herbicida com outras regiões da enzima, fora do sítio ativo.

Figura 36 - Perfil de força aplicado ao longo dos eixos x (preto), y (azul) e -y (vermelho).



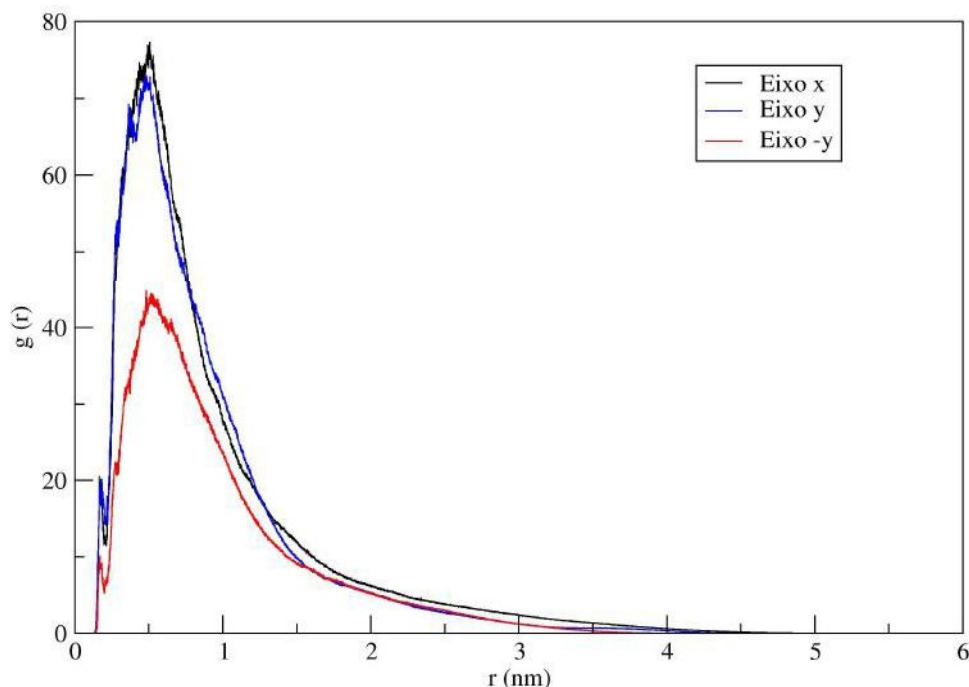
A análise da distribuição radial de pares entre o herbicida glifosato e o sítio ativo -S3P, proporciona a intensidade de interação entre eles e é definida como a razão entre o valor

médio da densidade numérica da molécula GPJ em uma casca esférica de raio r ao redor das partículas do sítio ativo-S3P e a densidade da molécula GPJ dentro de uma esfera de raio $r_{máx}$ (em geral, sendo a metade da aresta da caixa de simulação) ao redor do sítio ativo-S3P,

$$g_{sítio-S3P,GPJ}(r) = \frac{\langle \rho_{GPJ}(r) \rangle}{\langle \rho_{GPJ} \rangle_{local}} \quad (76)$$

Na Figura 37, a distribuição radial de pares é mostrada para os três eixos simulados via SMD. Nas três distribuições, as funções abrangem uma região de interação ampla (grande variação do raio r - eixo x) devido à região do sítio ativo-S3P ser relativamente grande, uma vez que é composta pelo substrato e por 19 resíduos de aminoácidos da enzima EPSPs. As funções atingem seus pontos máximos em torno de 0,5 a 0,55 nm, indicando que a densidade da molécula GPJ na casca esférica nesse raio é mais elevada do que a densidade do herbicida no sistema. A distribuição para os eixos x e y é bem próxima, uma vez que as forças das interações do herbicida no sítio ativo são quantitativamente próximas ao longo destes dois eixos, como visto no gráfico de perfil de força. No entanto, a altura do pico na simulação no eixo -y é acentuadamente menor que nos outros eixos, o que sugere uma interação menos efetiva entre o GPJ e o sítio ativo-S3P. Isso leva à hipótese de que o caminho de retirada do sítio ativo-S3P percorrido pelo herbicida ao longo do eixo -y sofre um menor impedimento em relação aos outros caminhos, consentindo com os resultados do perfil de força.

Figura 37 - Distribuição radial de pares nos três eixos: Eixo x (preto); Eixo y (azul) e Eixo -y (vermelho).

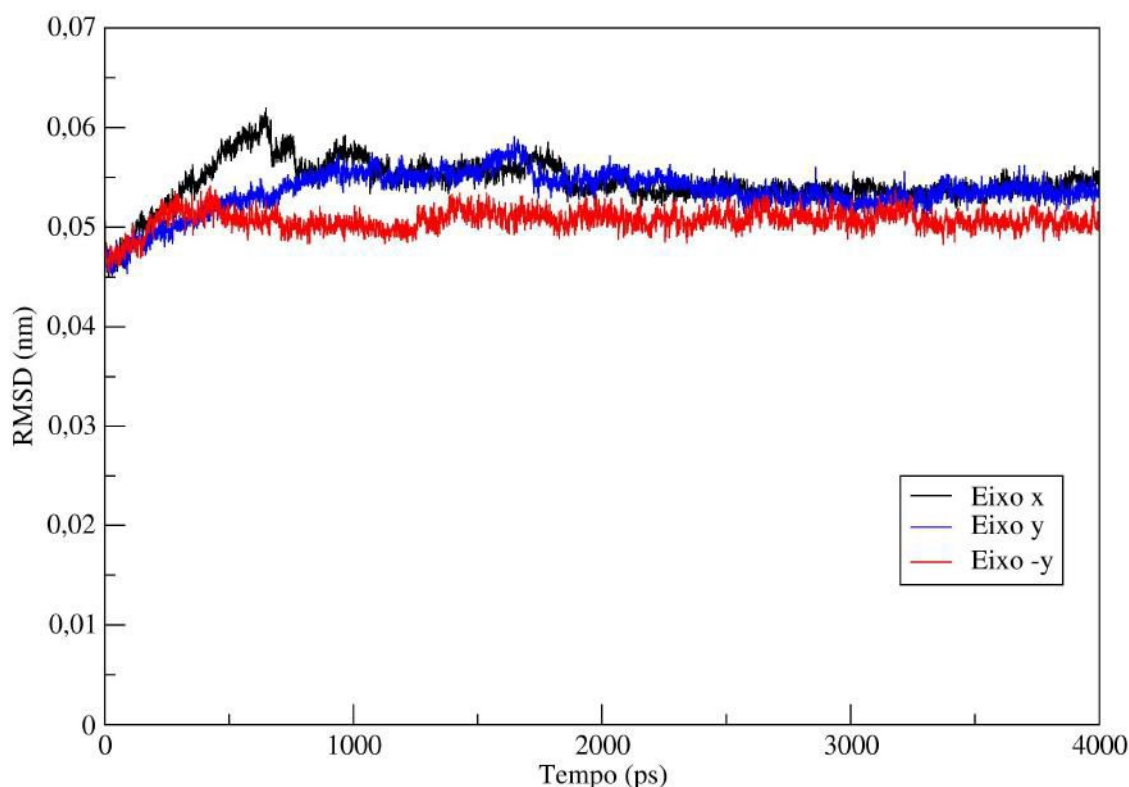


O RMSD da enzima EPSPs para as simulações ao longo dos três diferentes caminhos é mostrado na Figura 38. O desvio foi de 0,054 nm ($\pm 7,1 \cdot 10^{-4}$ nm), 0,054 nm ($\pm 9,9 \cdot 10^{-4}$ nm), 0,051 nm ($\pm 8,7 \cdot 10^{-4}$ nm) para os eixos x, y e -y respectivamente. O RMSD da enzima na dinâmica molecular direcional ao longo do eixo -y apresentou a menor flutuação, o que corrobora com os dados do perfil força e da distribuição radial de pares. Isto pode ser atribuído ao menor impedimento estérico ao longo desse caminho e consequentemente, uma menor interação do herbicida com a enzima. Fica evidente que, o caminho mais provável de desvinculação do herbicida em relação ao sítio ativo ocorre ao longo do eixo -y. No entanto, as simulações ao longo do eixo x, apresentaram um único pico de força bem definido, relacionado com as interações específicas do glifosato com o sítio ativo-S3P, que para a análise no AFM se torna importante. Apesar do caminho ao longo do eixo y apresentar um impedimento estérico maior em relação à retirada do herbicida do sítio ativo-S3P, o perfil em que aparece os maiores picos de força, ainda que, mais tardiamente ao longo do tempo, estão relacionados com as interações entre o glifosato e o sítio ativo.

Uma vez que, o mapeamento do potencial eletrostático ao longo de toda a superfície da enzima (seção 5.1) se mostrou bastante uniforme, não há como orientar, a princípio, uma funcionalização adequada da ponta do cantiléver do AFM de forma que se possa expor uma região específica do sítio ativo da enzima (como as que foram simuladas – x, y, -y) para

interagir com o herbicida glifosato. Isto posto, a funcionalização das enzimas sobre a ponta do cantiléver do AFM ocorrerá de forma aleatória, podendo esta apresentar diferentes orientações e consequentemente exposições diversas do sítio ativo para interagir com o glifosato.

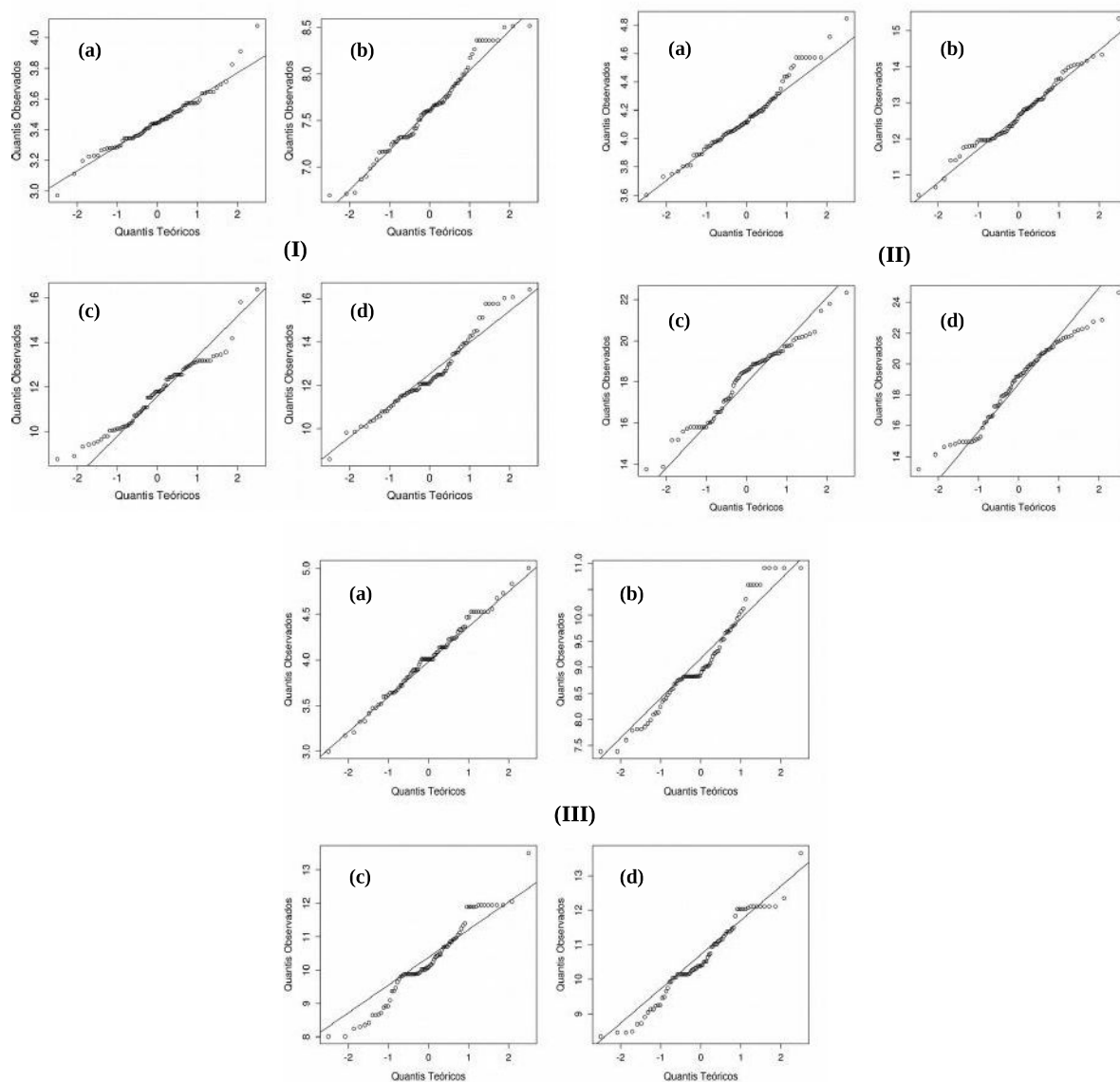
Figura 38 - Variação do RMSD da enzima EPSPs nas simulações de dinâmica molecular direcional ao longo dos eixos x (preto), y (azul) e -y (vermelho).



Para avaliar a normalidade de uma variável de forma qualitativa pode-se utilizar o gráfico Quantil-Quantil (Figura 39). Os resultados do teste não-paramétrico de Shapiro-Wilk dão resultados mais objetivos de aderência à distribuição Normal e são mostrados na Tabela 6. Nesse teste foram escolhidos 4 pontos ao longo da coordenada de reação (tempo de simulação), para cada eixo, para avaliar se a distribuição dos dados obtidos para o trabalho realizado em cada trajetória segue uma distribuição gaussiana. Foi escolhido um nível de significância de 10%, em que considerou-se o teste de hipótese como: H_0 : A distribuição do conjunto de dados segue uma distribuição normal e H_1 : A distribuição do conjunto de dados não segue uma distribuição normal.

Os Q-Q plots para os três eixos (Figura 39) apresentaram uma distribuição dos pontos próxima à reta indicando uma variável com distribuição normal.

Figura 39 - Gráfico Quantil-Quantil (I) – Eixo x, (a) trabalho a 300ps, (b) trabalho a 500ps, (c) trabalho a 1000ps, (d) trabalho a 1500ps; (II) – Eixo y, (a) trabalho a 500ps, (b) trabalho a 1000ps, (c) trabalho a 1500ps, (d) trabalho a 2000ps; (III) – Eixo -y, (a) trabalho a 300ps, (b) trabalho a 1000ps, (c) trabalho a 1500ps, (d) trabalho a 1750ps;



Pela Tabela 6, observa-se que todos os valores obtidos para o p-valor são maiores que 10%, logo, não se rejeita a hipótese zero, e desta forma a distribuição dos valores obtidos para o trabalho nos três eixos, nestes quatro pontos escolhidos ao longo da coordenada de reação, seguem uma distribuição gaussiana. Sendo assim, o número de simulações SMD realizadas é suficiente para a realização do cálculo do potencial de força média, segundo a metodologia da expansão cumulante da equação de Jarzynski.

Tabela 6 - Teste de Shapiro-Wilk

Eixos	P-Valor			
Eixo x	W em 300ps	W em 500ps	W em 1000ps	W em 1500ps
	0.02167	0.0863	0.02589	0.01948
Eixo y	W em 500ps	W em 1000ps	W em 1500ps	W em 2000ps
	0.09391	0.2919	0.0393	0.03267
Eixo -y	W em 300ps	W em 1000ps	W em 1500ps	W em 1750ps
	0.9526	0.01072	0.03206	0.05707

Desta forma, o Potencial de Força Média foi calculado a partir das simulações de SMD ao longo dos três eixos simulados. O PMF determina o perfil de energia livre ao longo da coordenada de reação, e foi calculado através da expansão cumulante de segunda ordem da igualdade de Jarzynski, equação 73.

A Figura 40 apresenta o PMF calculado para os três eixos ao longo do tempo de simulação. A menor barreira de energia livre ocorre ao longo do eixo -y, aproximadamente 10,7 kJ mol⁻¹, seguido pelo eixo x e eixo y, com barreiras de energia em torno de, 14,7 e 19,5 kJ mol⁻¹, respectivamente. Isso comprova os dados anteriores, em que, o processo de retirada do herbicida glifosato do sítio ativo da EPSPs ocorre favoravelmente ao longo do eixo -y.

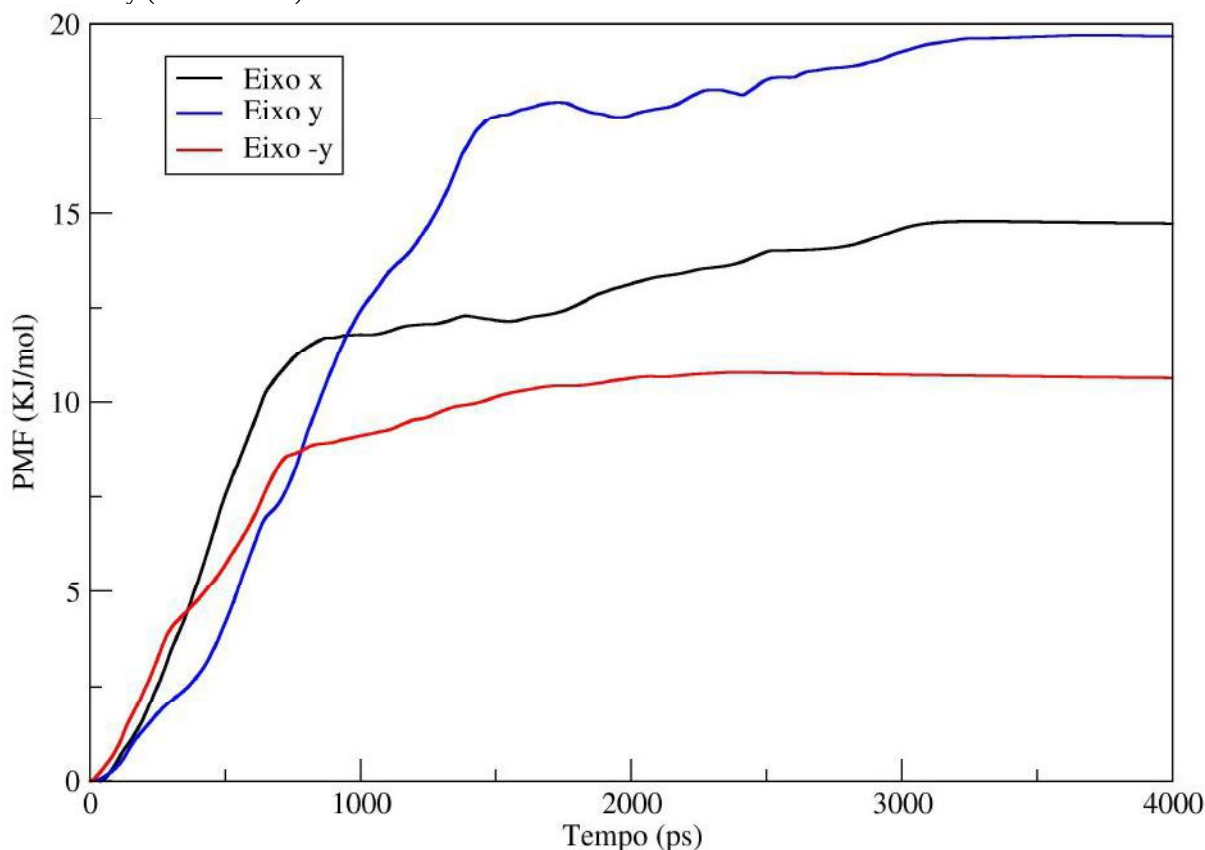
Esses valores de barreira de energia (processo de desacoplamento) podem ser comparados com a energia livre de ligação ($\Delta G_{\text{ligação}}$) para o complexo ligante-enzima (EPSPs-GPJ). Desta forma, para validar os cálculos de PMF, o $\Delta G_{\text{ligação}}$ foi calculado usando valores experimentais da constante de dissociação (K_d) na qual variou de 12 mM a 0,15 μM ¹²⁶, e a equação 77¹²⁷,

$$\Delta G_{\text{ligação}} = RT \ln(K_d) \quad (77)$$

onde, R é a constante universal dos gases e T é a temperatura em Kelvin.

Assim, o valor obtido para o $\Delta G_{\text{ligação}}$ pela equação 77, variou de -11.0 a -38.9 kJ mol⁻¹, sugerindo que a estimativa teórica do potencial de força média para calcular a energia envolvida no processo de retirada do herbicida do sítio ativo da enzima foi apropriada.

Figura 40 - Potencial de Força Média das simulações de SMD ao longo dos três eixos: x (em preto); y(em azul) e -y (em vermelho)



No entanto, cabe ressaltar que os valores experimentais de K_d foram obtidos em solução aquosa, na presença de uma força iônica apropriada, e as simulações realizadas no presente trabalho descrevem o sistema no vácuo (*ensemble* NVT) para simular os experimentos de AFM. Considerando que essas diferenças podem afetar significativamente os valores da energia de ligação, as simulações de SMD na presença de moléculas de água devem ser realizadas para prever a energia de ligação específica entre complexos enzima-ligante.

A funcionalização das enzimas sobre a ponta do cantiléver do AFM ocorrerá de forma aleatória, podendo esta apresentar diferentes orientações e consequentemente exposições diversas do sítio ativo para interagir com o glifosato. Assim, a energia de interação específica entre o glifosato e a EPSPs, deve variar entre as energias obtidas pelo potencial de força média nos três eixos apresentados neste trabalho.

6 Conclusão

O presente trabalho apresenta estudos de dinâmica molecular do sistema enzima-inibidor, especificamente o sistema EPSPs-Glifosato, com o intuito de obter propriedades moleculares e fazer previsões que possam auxiliar a construção de um nanobiossensor utilizando Microscopia de Força Atômica, para uma detecção específica e seletiva do herbicida glifosato.

Os cálculos de dinâmica molecular foram realizados utilizando como campo de força o OPLS-AA. Uma vez que, os parâmetros relacionados aos ângulos de torção do herbicida glifosato não são encontrados neste campo de força, parametrizações foram realizadas utilizando cálculos de mecânica quântica, utilizando as funções de base HF/6-31G*. Os parâmetros obtidos foram calculados utilizando o método de mínimos quadrados para comparação com os resultados obtidos via mecânica quântica e a validação foi realizada através de simulações por dinâmica molecular. Os resultados mostraram boa concordância entre os ângulos obtidos para as curvas de energia potencial clássica e quântica.

Dados de RMSD, RMSF, energia livre de solvatação e análise de ligações de hidrogênio dos resíduos do sítio ativo mostram que a enzima EPSPs, na presença do seu substrato Shikimato-3-Fosfato apresenta uma maior estabilidade conformacional em solução aquosa quando comparado ao sistema contendo apenas a enzima, e possivelmente apresentará uma melhor resposta na detecção do herbicida glifosato, após a funcionalização da ponta do cantiléver do AFM.

A emulação dos experimentos de AFM através de simulações por dinâmica molecular direcional foram realizadas aplicando-se forças externas para a retirada do herbicida do sítio ativo em três possíveis caminhos onde o impedimento estérico é menor. A análise do perfil de força e do Potencial de Força Média revelaram que o caminho com menor impedimento foi ao longo do eixo -y. Além disso, os valores de energia livre de interação entre o GPJ e a EPSPs a partir das simulações de SMD apresentaram a mesma ordem de magnitude de valores calculados a partir das constantes de dissociação do sistema, obtidas experimentalmente. O potencial eletrostático da superfície da enzima apresentou uma heterogeneidade na distribuição de cargas, e uma funcionalização direcionada necessita de estudos mais refinados.

7 Bibliografia

1. Vieira, R. *Fundamentos de Bioquímica - Textos Didáticos*. (2003).
2. Nelson, D. L. & Cox, M. M. *Lehninger Princípios de Bioquímica*. (2002).
3. Stryer, L., Berg, J. M. & Tymoczky, J. L. *Bioquímica*. (2004).
4. Marzzoco, A. *Bioquímica Básica*. (1999).
5. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
6. Motta, V. T. *Bioquímica*. (2011).
7. Herrmann, K. M. & Weaver, L. M. The Shikimate Pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 473–503 (1999).
8. Priestman, M. a., Healy, M. L., Funke, T., Becker, A. & Schönbrunn, E. Molecular basis for the glyphosate-insensitivity of the reaction of 5-enolpyruvylshikimate 3-phosphate synthase with shikimate. *FEBS Lett.* **579**, 5773–5780 (2005).
9. Stallings, W. C. *et al.* Structure and topological symmetry of the glyphosate target 5-enolpyruvylshikimate-3-phosphate synthase: a distinctive protein fold. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5046–5050 (1991).
10. Schönbrunn, E. *et al.* Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 1376–1380 (2001).
11. Pelaez, V., da Silva, L. R. & Araújo, E. B. Regulation of pesticides: A comparative analysis. *Sci. Public Policy* **40**, 644–656 (2013).
12. IBGE. *Indicadores de desenvolvimento sustentável - Brasil 2012*. (2012).
13. Galli, A. J. B. & Montezuma, M. C. in (ed. ACADCOM) (2005).
14. Coutinho, C. F. B. & Mazo, L. H. Complexos metálicos com o herbicida glifosato: Revisão. *Quim. Nova* **28**, 1038–1045 (2005).
15. McDowell, L. M. *et al.* Rotational-echo double-resonance NMR-restrained model of the ternary complex of 5-enolpyruvylshikimate-3-phosphate synthase. *J. Biomol. NMR* **28**, 11–29 (2004).
16. Sammons, R. D., Gruys, K. J., Anderson, K. S., Johnson, K. a & Sikorski, J. a. Reevaluating glyphosate as a transition-state inhibitor of EPSP synthase: identification of an EPSP synthase.EPSP.glyphosate ternary complex. *Biochemistry* **34**, 6433–6440 (1995).
17. Priestman, M. A. *et al.* Interaction of phosphonate analogues of the tetrahedral reaction intermediate with 5-enolpyruvylshikimate-3-phosphate synthase in atomic detail. *Biochemistry* **44**, 3241–3248 (2005).

18. Kaliannan, P., Mohamed Naseer Ali, M., Seethalakshmi, T. & Venuvanalingam, P. Electronic structure and conformation of glyphosate: An ab initio MO study. *J. Mol. Struct. THEOCHEM* **618**, 117–125 (2002).
19. De Amarante Junior, O. P., Dos Santos, T. C. R., Brito, N. M. & Ribeiro, M. L. Glifosato: Propriedades, toxicidade, usos e legislação. *Quimica Nova* **25**, 589–593 (2002).
20. Peixoto, M. M., Bauerfeldt, G. F., Herbst, M. H., Pereira, M. S. & da Silva, C. O. Study of the Stepwise Deprotonation Reactions of Glyphosate and the Corresponding pKa Values in Aqueous Solution. *J. Phys. Chem. A* **119**, 5241–5249 (2015).
21. Liu, B. *et al.* Thermodynamic Study on the Protonation Reactions of Glyphosate in Aqueous Solution: Potentiometry, Calorimetry and NMR spectroscopy. *J. Phys. Chem. B* **120**, 2132–2137 (2016).
22. Martins Júnior, H. A. Estudo de determinação de resíduos de glifosato e ácido aminometilfosfônico (AMPA) em amostras de soja e água usando cromatografia líquida acoplada à espectrometria de massas em tandem com ionização por electrospray (LC-ESI/MS/MS). 121 (2005).
23. Toni, L. R. M., De Santana, H. & Zaia, D. a M. Adsorção de glifosato sobre solos e minerais. *Quim. Nova* **29**, 829–833 (2006).
24. Londres, F. *AGROTÓXICOS NO BRASIL um guia para ação em defesa da vida*. Vasa (2011).
25. Benachour, N. & Séralini, G. E. Glyphosate formulations induce apoptosis and necrosis in human umbilical, embryonic, and placental cells. *Chem. Res. Toxicol.* **22**, 97–105 (2009).
26. Paganelli, A., Gnazzo, V., Acosta, H., López, S. L. & Carrasco, A. E. Glyphosate-based herbicides produce teratogenic effects on vertebrates by impairing retinoic acid signaling. *Chem. Res. Toxicol.* **23**, 1586–1595 (2010).
27. Sandrini, J. Z. *et al.* Effects of glyphosate on cholinesterase activity of the mussel *Perna perna* and the fish *Danio rerio* and *Jenynsia multidentata*: In vitro studies. *Aquat. Toxicol.* **130–131**, 171–173 (2013).
28. Ma, J. & Li, X. Alteration in the cytokine levels and histopathological damage in common carp induced by glyphosate. *Chemosphere* **128**, 293–298 (2015).
29. Chorfa, A. *et al.* Specific pesticide-dependent increases in α -synuclein levels in human neuroblastoma (SH-SY5Y) and melanoma (SK-MEL-2) cell lines. *Toxicol. Sci.* **133**, 289–297 (2013).
30. Gasnier, C. *et al.* Glyphosate-based herbicides are toxic and endocrine disruptors in human cell lines. *Toxicology* **262**, 184–191 (2009).
31. Mink, P. J., Mandel, J. S., Scurman, B. K. & Lundin, J. I. Epidemiologic studies of

- glyphosate and cancer: A review. *Regul. Toxicol. Pharmacol.* **63**, 440–452 (2012).
32. George, J., Prasad, S., Mahmood, Z. & Shukla, Y. Studies on glyphosate-induced carcinogenicity in mouse skin: A proteomic approach. *J. Proteomics* **73**, 951–964 (2010).
 33. Thongprakaisang, S., Thiantanawat, A., Rangkadilok, N., Suriyo, T. & Satayavivad, J. Glyphosate induces human breast cancer cells growth via estrogen receptors. *Food Chem. Toxicol.* **59**, 129–136 (2013).
 34. ANVISA. Anvisa esclarece sobre carcinogenicidade de cinco substâncias comercializados no país. (2015). Available at: <http://portal.anvisa.gov.br/wps/content/anvisa+portal/anvisa/sala+de+imprensa/menu+-+noticias+anos/2015/anvisa+esclarece+sobre+carcinogenicidade+de+cinco+substancias+comercializados+no+pais>. (Accessed: 3rd August 2015)
 35. Guyton, K. Z. *et al.* Carcinogenicity of tetrachlorvinphos, parathion, malathion, diazinon, and glyphosate. *Lancet Oncol.* **16**, 490–491 (2015).
 36. Marques, C. V. V. C. O., Marques, P. R. B. de O. & Nunes, G. S. Biossensor amperométrico para detecção screening de inseticidas carbamatos em águas de abastecimento. 81–92 (2006).
 37. Botero-Coy, A. M., Ibáñez, M., Sancho, J. V. & Hernández, F. Improvements in the analytical methodology for the residue determination of the herbicide glyphosate in soils by liquid chromatography coupled to mass spectrometry. *J. Chromatogr. A* **1292**, 132–141 (2013).
 38. Chu, S.-P. *et al.* Up-and-down-shaker-assisted dispersive liquid-liquid microextraction coupled with gas chromatography-mass spectrometry for the determination of fungicides in wine. *Food Chem.* **185**, 377–382 (2015).
 39. Ben-Zur, R., Hake, H., Hassoon, S., Bulatov, V. & Schechter, I. Optical analytical methods for detection of pesticides. *Rev. Anal. Chem.* **30**, 123–139 (2011).
 40. Martínez Gil, P., Laguarda-Miro, N., Camino, J. S. & Peris, R. M. Glyphosate detection with ammonium nitrate and humic acids as potential interfering substances by pulsed voltammetry technique. *Talanta* **115**, 702–705 (2013).
 41. De Amarante, O. P., Dos Santos, T. C. R., Nunes, G. S. & Ribeiro, M. L. Breve revisão de métodos de determinação de resíduos do herbicida á cido 2,4-diclorofenoxiacético (2,4-D). *Quim. Nova* **26**, 223–229 (2003).
 42. Leite, F. L. *et al.* Sensor arrays to detect humic substances and Cu(II) in waters. *Synth. Met.* **159**, 2333–2337 (2009).
 43. Binnig, G. & Quate, C. F. Atomic Force Microscope. *Phys. Rev. Lett.* **56**, 930–933 (1986).
 44. Ferreira, A. A. P. & Yamanaka, H. Microscopia de força atômica aplicada em

- imunoensaios. *Quim. Nova* **29**, 137–142 (2006).
45. Herrmann, P. S. P. *et al.* Microscopia de varredura por força: uma ferramenta poderosa no estudo de polímeros. *Polímeros* **7**, 51–61 (1997).
 46. Howland, R. & Benatar, L. in 87 (1996).
 47. Timhadjelt, L., Serier, A., Belgacem, M. N. & Bras, J. Elaboration of cellulose based nanobiocomposite: Effect of cellulose nanocrystals surface treatment and interface ‘melting’. *Ind. Crops Prod.* **72**, 7–15 (2015).
 48. Bald, I. & Keller, A. Molecular Processes Studied at a Single-Molecule Level Using DNA Origami Nanostructures and Atomic Force Microscopy. *Molecules* **19**, 13803–13823 (2014).
 49. Köhler, M. *et al.* pH-Dependent Deformations of the Energy Landscape of Avidin-like Proteins Investigated by Single Molecule Force Spectroscopy. *Molecules* **19**, 12531–12546 (2014).
 50. Friedrichs, J. *et al.* A practical guide to quantify cell adhesion using single-cell force spectroscopy. *Methods* **60**, 169–178 (2013).
 51. Kashef, J. & Franz, C. M. Quantitative methods for analyzing cell–cell adhesion in development. *Dev. Biol.* **401**, 165–174 (2014).
 52. Franca, E. F., Leite, F. L., Cunha, R. A., Oliveira, O. N. & Freitas, L. C. G. Designing an enzyme-based nanobiosensor using molecular modeling techniques. *Phys. Chem. Chem. Phys.* **13**, 8894–8899 (2011).
 53. da Silva, A. C. N. *et al.* Nanobiosensors based on chemically modified AFM probes: A useful tool for metsulfuron-methyl detection. *Sensors (Switzerland)* **13**, 1477–1489 (2013).
 54. Lee, J. E. *et al.* A comparative study on antibody immobilization strategies onto solid surface. *Korean J. Chem. Eng.* **30**, 1934–1938 (2013).
 55. de Thier, P., Bacharouche, J., Duval, J. F. L., Skali-Lami, S. & Francius, G. Atomic force microscopy analysis of IgG films at hydrophobic surfaces: A promising method to probe IgG orientations and optimize ELISA tests performance. *Biochim. Biophys. Acta - Proteins Proteomics* **1854**, 138–145 (2015).
 56. de Oliveira, G. S. Análise da Interação molecular proteína-herbicida através de simulação computacional: Aplicação no desenvolvimento de nanobiossensores. (Universidade Federal de São Carlos, 2013).
 57. Etchegaray, A., Bueno, C. de C. & Teschke, O. Identificação de microcistina LR ao nível molecular empregando microscopia de força atômica. *Quim. Nova* **33**, 1843–1848 (2010).
 58. Ito, T., Ibrahim, S. & Grabowska, I. Chemical-force microscopy for materials

- characterization. *TrAC - Trends Anal. Chem.* **29**, 225–233 (2010).
59. Comba, P., Hambley, T. W. & Peter Comba, T. W. H. *Molecular Modeling of Inorganic Compounds*. (2001).
 60. Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. (2010). doi:10.1007/978-1-4419-6351-2
 61. Sant’Anna, C. M. R. Molecular modeling methods in the study and design of bioactive compounds: An introduction. *Rev. Virtual Química* **1**, (2009).
 62. Cramer, C. J. *Essentials of Computational Chemistry Theories and Models. Essentials of Computational Chemistry* **42**, (2004).
 63. Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations. Interpreting* (2003).
 64. dos Santos, P. S. Estudo dos mecanismos de formação das espécies PN e NS no meio interestelar, em cometas e atmosferas planetárias. (PUC-Rio, 2009).
 65. Trsic, M. & Pinto, M. F. S. *Química Quântica - Fundamentos e Aplicações*. (2009).
 66. Jensen, F. *Introduction to Computational Chemistry*. (John Wiley & Sons Ltd, 2007).
 67. Young, D. C. *Computational chemistry: A Practical Guide for Applying Techniques to Real-World Problems*. (John Wiley & Sons, Inc., 2001).
 68. Laschuk, E. F. Novo Formalismo Semi-Empírico para Cálculos Químico-Quânticos. (Universidade Federal do Rio Grande do Sul, 2005).
 69. Hinchliffe, A. *Molecular Modelling for Beginners*. (WILEY, 2008).
 70. Coelho, L. W., Junqueira, G. M. A., Machuca Herrera, J. O., Machado, S. D. P. & Machado, B. D. C. Aplicação de mecânica molecular em química inorgânica. *Quim. Nova* **22**, 396–404 (1999).
 71. Namba, a. M., Da Silva, V. B. & Da Silva, C. H. T. P. Dinâmica Molecular: Teoria e aplicações em planejamento de fármacos. *Eclét. Quim.* **33**, 13–24 (2008).
 72. Leach, A. R. *Molecular Modelling. Principles and Applications*. (2001).
 73. Morgon, N. H. & Coutinho, K. *Métodos de Química Teórica e Modelagem Molecular*. (2007).
 74. Stariolo, D. A. *Mecânica Estatística*. (Universidade Federal do Rio Grande do Sul, 2014).
 75. Censoni, L. B. Dinâmica molecular e redes complexas no estudo da difusão térmica em xilanases da família 11. (Universidade de São Paulo, 2013).
 76. Loup, V. Computer ‘Experiments’ on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **159**, 98–103 (1967).

77. Allen, M. P. & Tildesley, D. J. *Computer Simulations of Liquids*. (Oxford University, 1989).
78. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. . Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
79. Andersen, H. C. Rattle: A ‘velocity’ version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **52**, 24–34 (1983).
80. van Gunsteren, W. F. & Karplus, M. Effect of constraints on the dynamics of macromolecules. *Macromolecules* **15**, 1528–1544 (1982).
81. Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **64**, 253–287 (1921).
82. Tironi, I. G., Sperb, R., Smith, P. E. & van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451 (1995).
83. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089 (1993).
84. Fukuda, I. & Nakamura, H. Non-Ewald methods: Theory and applications to molecular systems. *Biophys. Rev.* **4**, 161–170 (2012).
85. Bizzarri, A. R. & Cannistraro, S. Molecular dynamics of water at the protein-solvent interface. *J. Phys. Chem. B* **106**, 6617–6633 (2002).
86. Makarov, V., Pettitt, B. M. & Feig, M. Solvation and hydration of proteins and nucleic acids: A theoretical view of simulation and experiment. *Acc. Chem. Res.* **35**, 376–384 (2002).
87. Nutt, D. R. & Smith, J. C. Molecular Dynamics Simulations of Proteins: Can the Explicit Water Model Be Varied? *J. Chem. Theory Comput.* **3**, 1550–1560 (2007).
88. Bashford, D. & Case, D. A. Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **51**, 129–152 (2000).
89. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).
90. Mahoney, M. W. & Jorgensen, W. L. Quantum, intramolecular flexibility, and polarizability effects on the reproduction of the density anomaly of liquid water by simple potential functions. *J. Chem. Phys.* **112**, 8910–8922 (2000).
91. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & Hermans, J. Interaction models for water in relation to protein hydration. *Intermol. Forces* 331–342 (1981).
92. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, a & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690

- (1984).
93. Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **52**, 255–268 (1984).
 94. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
 95. Bond, S. D., Leimkuhler, B. J. & Laird, B. B. The Nosé – Poincaré Method for Constant Temperature Molecular Dynamics. *J. Comput. Phys.* **151**, 114–134 (1999).
 96. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182 (1981).
 97. Debye, P. Näherungsformeln für die Zylinderfunktionen für große Werte des Arguments und unbeschränkt veränderliche Werte des Index. *Math. Ann.* **67**, 535–558 (1909).
 98. Hestenes, M. R. & Stiefel, E. Methods of Conjugate Gradients for Solving Linear Systems. *J. Res. Natl. Bur. Stand.* (1934). **49**, 409–436 (1952).
 99. Tjalling, J. Y. Historical development of the Newton-Raphson method. *SIAM Rev.* **37**, 531–551 (1995).
 100. Van Der Spoel, D. *et al.* GROMACS User Manual Version 4.5. *Computer Programs in* **18**, 370 (2010).
 101. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
 102. Hess, B., Kutzner, C., Van Der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
 103. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and Testing of the OLPS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
 104. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
 105. Stewart, J. P. J. MOPAC 2009. *Stewart Computational* Available at: [HTTP://openmopac.net](http://openmopac.net).
 106. Nesse, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 73–78 (2012).
 107. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollman, P. a. Application of RESP Charges To Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **115**, 9620–9631 (1993).
 108. Woods, R. J. & Chappelle, R. Restrained electrostatic potential atomic partial charges

- for condensed-phase simulations of carbohydrates. *J. Mol. Struct. THEOCHEM* **527**, 149–156 (2000).
109. Valiev, M. *et al.* NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **181**, 1477–1489 (2010).
 110. Humphrey, W., Dalke, A. & Schulten, K. VMD - Visual Molecular Dynamics. *J. Molec. Graph.* **14**, 33–38 (1996).
 111. Holst, M., Baker, N. & Wang, F. Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. Algorithms and examples. *J. Comput. Chem.* **21**, 1319–1342 (2000).
 112. Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y. & Schulten, K. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72**, 1568–1581 (1997).
 113. Park, S. & Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **120**, 5946–5961 (2004).
 114. Zhang, J.-L., Zheng, Q.-C., Li, Z.-Q. & Zhang, H.-X. Molecular dynamics simulations suggest ligand's binding to nicotinamidase/pyrazinamidase. *PLoS One* **7**, e39546 (2012).
 115. Amarante, A. M. *et al.* Modeling the coverage of an AFM tip by enzymes and its application in nanobiosensors. *J. Mol. Graph. Model.* **53**, 100–104 (2014).
 116. Lemkul, J. A. & Bevan, D. R. Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics. *J. Phys. Chem. B* **114**, 1652–1660 (2010).
 117. Franca, Eduardo de Faria, Leite, Fábio L. & Amarante, F. L. in *MICROSCOPY: SCIENCE, TECHNOLOGY, APPLICATIONS AND EDUCATION* 1338–1349 (2010).
 118. Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **56**, 5018–5035 (1997).
 119. Perišić, O. & Lu, H. On the Improvement of Free-Energy Calculation from Steered Molecular Dynamics Simulations Using Adaptive Stochastic Perturbation Protocols. *PLoS One* **9**, e101810 (2014).
 120. Hummer, G. Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.* **114**, 7330–7337 (2001).
 121. Park, S., Khalili-Araghi, F., Tajkhorshid, E. & Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* **119**, 3559 (2003).
 122. Eaton, J. W., Bateman, D. & Hauberg, S. GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations. (2009).
 123. Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).

124. Team, R. C. R: A language and environment for statistical computing. (2013).
125. Apweiler, R. & Birney, E. The European Bioinformatics Institute. Available at: <https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html>. (Accessed: 1st January 2016)
126. Ream, J. E., Yuen, H. K., Frazier, R. B. & Sikorski, J. A. EPSP synthase: Binding studies using isothermal titration microcalorimetry and equilibrium dialysis and their implications for ligand recognition and kinetic mechanism. *Biochemistry* **31**, 5528–5534 (1992).
127. Copeland, R. A. *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*. (Wiley-VCH, Inc., 2000).